

University of Groningen

Spatial release from informational masking declines with age

Zobel, Benjamin H.; Wagner, Anita; Sanders, Lisa D.; Baskent, Deniz

Published in:
Journal of the Acoustical Society of America

DOI:
[10.1121/1.5118240](https://doi.org/10.1121/1.5118240)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Zobel, B. H., Wagner, A., Sanders, L. D., & Baskent, D. (2019). Spatial release from informational masking declines with age: Evidence from a detection task in a virtual separation paradigm. *Journal of the Acoustical Society of America*, 146(1), 548-566. <https://doi.org/10.1121/1.5118240>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Spatial release from informational masking declines with age: Evidence from a detection task in a virtual separation paradigm

Benjamin H. Zobel, Anita Wagner, Lisa D. Sanders, and Deniz Başkent

Citation: *The Journal of the Acoustical Society of America* **146**, 548 (2019); doi: 10.1121/1.5118240

View online: <https://doi.org/10.1121/1.5118240>

View Table of Contents: <https://asa.scitation.org/toc/jas/146/1>

Published by the *Acoustical Society of America*

ARTICLES YOU MAY BE INTERESTED IN

[Observations of phase and intensity fluctuations for low-frequency, long-range transmissions in the Philippine Sea and comparisons to path-integral theory](#)

The Journal of the Acoustical Society of America **146**, 567 (2019); <https://doi.org/10.1121/1.5118252>

[Cognitive factors contribute to speech perception in cochlear-implant users and age-matched normal-hearing listeners under vocoded conditions](#)

The Journal of the Acoustical Society of America **146**, 195 (2019); <https://doi.org/10.1121/1.5116009>

[Intelligibility of conversational and clear speech in young and older talkers as perceived by young and older listeners](#)

The Journal of the Acoustical Society of America **146**, EL28 (2019); <https://doi.org/10.1121/1.5116322>

[Compensation of array lens effects for improved co-registration of passive acoustic mapping and B-mode images for cavitation monitoring](#)

The Journal of the Acoustical Society of America **146**, EL78 (2019); <https://doi.org/10.1121/1.5118238>

[Impact of low-pass filtering on masking release in understanding Mandarin sentences](#)

The Journal of the Acoustical Society of America **146**, EL52 (2019); <https://doi.org/10.1121/1.5117173>

[Deep metric learning for bioacoustic classification: Overcoming training data scarcity using dynamic triplet loss](#)

The Journal of the Acoustical Society of America **146**, 534 (2019); <https://doi.org/10.1121/1.5118245>



WHY PUBLISH WITH US?

Spatial release from informational masking declines with age: Evidence from a detection task in a virtual separation paradigm

Benjamin H. Zobel^{a)}

Department of Psychological and Brain Sciences, University of Massachusetts, Amherst, Massachusetts 01003, USA

Anita Wagner^{b)}

Department of Otorhinolaryngology-Head and Neck Surgery, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

Lisa D. Sanders

Department of Psychological and Brain Sciences, University of Massachusetts, Amherst, Massachusetts 01003, USA

Deniz Başkent^{b)}

Department of Otorhinolaryngology-Head and Neck Surgery, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

(Received 19 December 2018; revised 25 June 2019; accepted 28 June 2019; published online 29 July 2019)

Declines in spatial release from informational masking may contribute to the speech-processing difficulties that older adults often experience within complex listening environments. The present study sought to answer two fundamental questions: (1) Does spatial release from informational masking decline with age and, if so, (2) does age predict this decline independently of age-typical hearing loss? Younger (18–34 years) and older (60–80 years) adults with age-typical hearing completed a yes/no target-detection task with low-pass filtered noise-vocoded speech designed to reduce non-spatial segregation cues and control for hearing loss. Participants detected a target voice among two-talker masking babble while a virtual spatial separation paradigm [Freyman, Helfer, McCall, and Clifton, *J. Acoust. Soc. Am.* **106**(6), 3578–3588 (1999)] was used to isolate informational masking release. The younger and older adults both exhibited spatial release from informational masking, but masking release was reduced among the older adults. Furthermore, age predicted this decline controlling for hearing loss, while there was no indication that hearing loss played a role. These findings provide evidence that declines specific to aging limit spatial release from informational masking under challenging listening conditions. © 2019 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.1121/1.5118240>

[VB]

Pages: 548–566

I. INTRODUCTION

Declines in auditory processing at peripheral and central levels can result in significant communication problems in many older adults. Challenging multi-talker listening conditions, common in everyday life, require successful encoding, localization, segregation, and selective processing of speech signals. These are precisely the conditions in which older adults often experience the greatest listening difficulties (for reviews, see CHABA, 1988; Gordon-Salant, 2005; Pichora-Fuller *et al.*, 2017). In typical listening situations, it is rarely, if ever, the case that the source of a relevant speech signal (target) and the source of competing noise (masker) are perfectly co-located in space. Therefore, a common form of the cocktail party problem (Cherry, 1953) is the challenge of preferentially processing target speech that originates at one

location relative to masking speech that originates at a separate location. As such, the benefit to speech processing that is realized when target and masker are spatially separated compared to spatially co-located—a phenomenon known as spatial release from masking (for reviews, see Bronkhorst, 2000, 2015)—reflects an important aspect of successful speech processing in everyday life. It follows that any declines in spatial release from masking that may be experienced in older age could result in greater difficulty understanding speech within complex multi-talker environments.

Studies of spatial release from masking must consider the stages of auditory processing involved in successful listening under different conditions. An important distinction has been made between energetic masking and informational masking. Energetic masking describes the sensory masking that occurs under conditions in which the masker energy dominates the target energy. That is, a poor signal-to-noise ratio (SNR; ratio of target intensity relative to masker intensity) distributed across the spectral region of the target will prevent target encoding (Fletcher, 1940; Miller, 1947). Zurek (1993) showed that under simple listening conditions,

^{a)}Electronic mail: bzobel@psych.umass.edu

^{b)}Also at: Graduate School of Medical Sciences, Research School of Behavioral and Cognitive Neurosciences, University of Groningen, Groningen, the Netherlands.

much of the benefit of spatial separation can be explained by a release from energetic masking through head shadow and binaural interaction (also see introduction in [Freyman et al., 1999](#)). That is, spatial separation can allow the head to cast a beneficial acoustic shadow on the ear that is farther from the masker, partially attenuating the masker energy relative to the target energy at this better ear, particularly across the higher frequencies ([Shaw, 1974](#)). In addition, spatial separation can introduce interaural timing differences such that binaural interactions within the auditory system partially release energetic masking, particularly across the lower frequencies. This binaural interaction is exemplified by the improvement in the processing of a target in broadband noise when either the target or the noise is presented out of phase interaurally compared to both target and noise being presented in phase, a phenomenon commonly referred to as the binaural masking level difference (BMLD) ([Hirsh, 1948](#); [Licklider, 1948](#)).

Informational masking can be described as any masking that is not accounted for by energetic masking and arises from target/masker confusion (for review, see [Kidd et al., 2008](#)). Complex multi-talker conditions can involve a substantial amount of informational masking ([Carhart et al., 1969](#); [Freyman et al., 1999](#)). However, the effects of informational masking are also evident in the detection of far simpler targets (e.g., [Durlach et al., 2003](#); [Kidd et al., 1994](#); [Lutfi, 1993](#); [Lutfi et al., 2013](#); [Neff, 1995](#); [Neff and Green, 1987](#); [Oh and Lutfi, 2000](#); [Watson et al., 1976](#); [Watson et al., 1975](#)). For example, [Kidd et al. \(1994\)](#) asked listeners to detect a target tone when simultaneously presented with a multi-tone masker. Although energetic masking was consistently low across all conditions, because none of the frequencies in the masker complex fell within a protected spectral region around the target, informational masking was high under conditions in which listeners had difficulty perceptually separating the target tone from the combined target-masker complex. Release from informational masking was observed under conditions in which differences between the target and masker, including perceived spatial separation, allowed for perceptual separation of the two. In the present study, informational masking was maximized by making maskers and targets perceptually similar, and by making the content of maskers and targets and the timing of targets unpredictable ([Lutfi et al., 2013](#)).

Age-related changes in spatial release from masking have been investigated, though not in a manner that has systematically distinguished the effects of aging from those of hearing loss and informational masking release from energetic masking release (for review, see [Glyde et al., 2011](#)). Some studies report no age-specific declines in spatial release from masking ([Füllgrabe et al., 2015](#); [Glyde et al., 2013](#); [Jakien and Gallun, 2018](#); [Jakien et al., 2017](#)), while others claim to demonstrate them (e.g., [Gallun et al., 2013](#)). Still other studies show age as the dominant predictor of declines under some conditions and hearing loss as the dominant predictor under other conditions (e.g., [Srinivasan et al., 2016](#); [Srinivasan et al., 2017](#)). All of these studies have used actual or simulated (i.e., head-related transfer functions) physical separation between target and masker, which can

introduce head shadow and binaural interaction effects that make it difficult to isolate effects associated with informational masking release from those associated with energetic masking release, even when maskers are symmetrically separated ([Kidd et al., 2010](#)). Inconsistencies among these studies may arise from differential influences of age and hearing loss on informational and energetic masking release that cannot be easily resolved under conditions of physical separation. Binaural interaction may particularly confound the ability to isolate age-specific effects of informational masking release since older adults with clinically normal hearing have shown reduced BMLDs ([Anderson et al., 2018](#); [Eddins and Eddins, 2018](#); [Eddins et al., 2018](#); [Grose et al., 1994](#); [Pichora-Fuller and Schneider, 1991](#)) and BMLDs may be reduced by even slight hearing loss ([Bernstein and Trahiotis, 2018](#)). In the present study, release from informational masking was achieved with a specific type of spatial separation that minimizes differences in energetic masking across spatial conditions.

[Freyman et al. \(1999\)](#) employed such a *virtual separation* paradigm that avoids the conflation of energetic and informational masking release inherent to paradigms that manipulate physical spatial separation. Specifically, listeners are positioned with one loudspeaker directly in front of them and another loudspeaker to their right. In the spatially collocated condition, masking speech and target speech are both presented from the front loudspeaker while no stimulus is presented from the right loudspeaker (F-F condition; the first “F” refers to the location of the target and the second to the location of the masker). In the virtual spatially separated condition, target and masker are again presented from the front loudspeaker while an identical copy of the masker is presented from the right loudspeaker such that the onset of the right masker precedes the onset of the front masker by 4 ms (F-RF condition, referring to the front location of the target and the right-leading-front locations of the maskers). The 4-ms stimulus onset asynchrony (SOA) between the identical maskers creates the precedence effect (for review, see [Litovsky et al., 1999](#)), resulting in the perception of only a single masker coming from the right. Thus, in the F-RF condition, listeners hear a target from the front and a masker from the right (virtual spatial separation), despite the fact that the front loudspeaker presents both target and masker (no physical spatial separation).

Multiple studies have demonstrated that the substantial release from masking observed in the F-RF condition cannot be explained as reduction in energetic masking ([Brungart et al., 2005](#); [Freyman et al., 2001, 2004, 2008](#); [Freyman et al., 1999](#); [Morse-Fortier et al., 2017](#); [Rakerd et al., 2006](#)). Specifically, masking release is only observed in the F-RF condition to the extent that informational masking is present in the F-F condition. When the potential for energetic masking is high because the masker spectrally overlaps the target at a constant intensity, but the potential for informational masking is low because the target (e.g., speech) and masker (e.g., steady-state broadband noise) are perceptually distinct, little to no benefit is observed in the F-RF condition for the detection or identification of natural or vocoded speech (some energetic masking effects have been reported, but

only to varying degrees at SOAs ≤ 2 ms; Brungart *et al.*, 2005; Freyman *et al.*, 2001, 2004, 2008, 1999; Morse-Fortier *et al.*, 2017; Rakerd *et al.*, 2006). Freyman *et al.* (1999) directly tested release from energetic masking under conditions of physical and virtual separation. Listeners were asked to detect narrowband noise-burst targets (1/3-octave bandwidths centered at 250–6300 Hz) in steady-state broadband speech-shaped noise. Participants exhibited ~ 9 dB of masking release across the range of frequency-centered targets when targets were physically separated from maskers (F-R condition), consistent with the release from energetic masking by head shadow and binaural interaction predicted by the model of Zurek (1993). In contrast, when targets and maskers were virtually separated in the F-RF condition, no appreciable change in energetic masking was found. As such, any benefits observed in the F-RF condition may be attributed primarily to release from the informational masking that is present in the F-F condition. Thus, differences in the extent to which younger and older adults benefit from virtual spatial separation provides evidence for age-related changes in spatial release from informational masking.

A handful of studies have examined spatial release from informational masking in younger and older adults using virtual separation (Avivi-Reich *et al.*, 2014; Helfer *et al.*, 2010; Helfer and Freyman, 2008; Li *et al.*, 2004). However, the extent to which—or even if—spatial release from informational masking declines with age remains unclear. For example, Li *et al.* (2004) measured younger and older adults' ability to identify words in target sentences presented with two-talker babble. In addition to presenting target and babble in spatially co-located, and virtually separated (3-ms SOA) conditions, the researchers also included a condition in which target and babble were virtually separated using a 0-ms SOA between the babble presentations. The latter condition, referred to here as the F-SUM condition, elicits summing localization (Blauert, 1997) to create the perception of a single masking babble located between the two loudspeakers. No substantive difference in spatial release from informational masking was found between the age groups; psychometric functions for younger and older adults were essentially identical to each other in all spatial conditions after applying a simple correction of 2.8 dB SNR for the older adults. In contrast, other studies using the virtual separation paradigm have reported some degree of age-related reduction in spatial release from informational masking (Helfer *et al.*, 2010; Helfer and Freyman, 2008). However, older adults tended to exhibit poorer performance in all conditions, making it difficult to compare the size of spatial effects between the age groups. Avivi-Reich *et al.* (2014) showed similar performance of younger and older adults in the F-F condition, and a small reduction (~ 2 dB SNR) in spatial release from masking for the older adults. However, masking release was uncommonly small for both age groups, likely because the 12-talker masking babble used for the R-SPIN test was perceptually distinct enough from the target that there was only a small amount of informational masking in the F-F condition that could be released in the F-RF condition (Freyman *et al.*, 2004).

As previously discussed, confounds between informational and energetic masking release in studies that have used physical separation between targets and maskers have not allowed for an assessment of the independent contributions of age and hearing loss on spatial release from informational masking. Even in the studies that have isolated informational masking release using virtual separation, however, the independent effects of age remain unclear. Older adults in these studies had audiometric thresholds that generally did not exceed a categorization of “mild” hearing loss [25–40 dB hearing level (HL); Clark, 1981] below 4000 Hz while greater hearing loss was exhibited at higher frequencies. Inclusion criteria ranged from more conservative (< 25 dB HL at 250–3000 Hz; Avivi-Reich *et al.*, 2014; Li *et al.*, 2004) to more liberal (≤ 30 dB HL at 250–2000 Hz; Helfer *et al.*, 2010; Helfer and Freyman, 2008). However, the extent to which age predicted effects, independent of the hearing differences between the younger and older adults, was not measured in a manner that provided conclusive evidence. The present study was designed to describe the psychometric functions for younger and older adults in each spatial condition and maximize statistical power for detecting age effects on spatial release from informational masking while thoroughly controlling for the inevitable differences in hearing.

Because of the lack of clarity from the results described above, the present study sought to answer two open questions: (1) Does spatial release from informational masking—i.e., reduction in target/masker confusion afforded by different localizations of target and masker—decline with age and, if so, (2) does age predict this decline independently of age-typical hearing loss? To answer these questions, younger and older adults with age-typical hearing were tested with the virtual separation paradigm with three important modifications. First, the present study used noise-vocoded speech (Freyman *et al.*, 2008; Qin and Oxenham, 2003) rather than natural speech. To isolate the effects of spatial release from informational masking and make certain that any difference in age groups was driven by differences in the use of those spatial cues, it was important to minimize non-spatial differences in targets and maskers. Natural speech targets and maskers typically differ in voice pitch, timbre, prosody, linguistic content, and the extent to which they have been primed, all of which can facilitate release from informational masking (e.g., Başkent and Gaudrain, 2016; Bradlow and Alexander, 2007; Brungart, 2001; Culling and Summerfield, 1995; Darwin and Hukin, 2000; Darwin *et al.*, 2003; El Boghdady *et al.*, 2019; Freyman *et al.*, 2001, 2004; Freyman *et al.*, 2005; Mattys *et al.*, 2012; Vestergaard *et al.*, 2009). The non-spatial differences between natural speech targets and maskers result in thresholds that are ~ 3 –6 dB SNR lower in the F-F condition than what is observed for vocoded speech (Freyman *et al.*, 2008; Morse-Fortier *et al.*, 2017). Further, the availability of multiple cues to distinguish natural speech targets and maskers results in greater variability in masking thresholds both across and within individuals (Freyman *et al.*, 2008; Morse-Fortier *et al.*, 2017). Thus, the use of vocoded stimuli in the present study was expected to provide several advantages. By reducing non-spatial cues, any group differences in release from informational masking

could be attributed to the spatial cues. In addition, any age-related effects of non-spatial cues would be reduced, making it more likely to observe similar performance in younger and older adults in the baseline F-F condition, allowing for a comparable measure of spatial release from informational masking across age groups. Finally, reductions in variability would increase the sensitivity of comparisons between groups and across spatial conditions.

The second notable feature of the present study was the use of a detection task rather than a speech identification task (e.g., word discrimination, recognition, or comprehension). Performance on speech-identification tasks is more likely to be influenced by language experience and proficiency (e.g., vocabulary size) and the ability to manage multiple cognitive demands (e.g., tracking and holding words in working memory). By reducing task demands, a detection task should reduce differences in performance of younger and older adults that are unrelated to spatial release from informational masking. Moreover, prior research has shown larger differences in thresholds for the F-F and F-RF conditions when using detection compared to speech-identification tasks (Freyman *et al.*, 2008). Greater sensitivity in the measure of spatial release from informational masking was expected to provide greater power to detect age-related effects.

The third distinguishing feature of the present study was the application of a low-pass filter to the stimuli in an effort to control for hearing loss. A sharp 2-kHz cutoff was chosen to simulate profound high-frequency hearing loss across all participants. This control was intended to minimize any disadvantage of high-frequency hearing loss typical among older adults (Hannula *et al.*, 2011; Hoffman *et al.*, 2012) while still producing stimuli that would elicit strong spatial release from informational masking, as supported by pilot data. Incorporating a hearing-loss control into the stimuli itself has been shown to provide greater power to tease apart independent effects of age and hearing loss on spatial release from masking (Gallun *et al.*, 2013). In addition to increasing the power to isolate age-specific effects, the control was also intended to better match the performance of younger and older adults in the baseline F-F condition.

The present study design allowed for detection thresholds for younger and older adults to be estimated based on psychometric functions fit to detection rates collected in the F-F, F-RF, F-SUM conditions. F-SUM was included to determine whether any differences observed in the F-RF condition were specific to the precedence effect. Performance in the baseline F-F condition was predicted to be similar for younger and older adults, such that age-related declines in spatial release from informational masking could be clearly assessed across spatial conditions. An observed reduction in masking release among the older adults would motivate statistical analyses to assess whether age and/or hearing loss (based on pure-tone audiometry) independently predicted the decline.

II. METHODS

A. Participants

Twenty-two younger adults (15 female, range = 18–34 years, $M = 22.41$ years, $SD = 4.34$ years) and 22 older adults

(10 female, range = 60–80 years, $M = 67.45$ years, $SD = 6.04$ years) contributed data for analysis.¹ Participants were native Dutch speakers reporting no diagnosed hearing problems or neurological disorders and no use of psychoactive medication at the time of the study. The Mini-Mental State Examination (Folstein *et al.*, 1975) was administered to the older participants and did not indicate abnormalities in cognitive function (all scores $\geq 28/30$). Data were excluded from three additional older adults who failed to complete the study due to fatigue ($n = 1$), difficulty understanding instructions ($n = 1$), and self-reported hearing problems with an audiogram showing “moderate” (41–55 dB HL; Clark, 1981) hearing loss below 2000 Hz ($n = 1$). All procedures were conducted in accordance with the review and approval of the Medical Ethical Committee of the University Medical Center Groningen. Participants provided verbal and written consent prior to beginning the study and were compensated at a rate of €8 per hour of participation plus travel expenses, in accordance with departmental policy.

B. Hearing assessments

Hearing was assessed with pure-tone air-conduction audiometry performed in each ear. Figure 1 shows the hearing thresholds measured at 250–8000 Hz for all participants. Thresholds for the younger participants were ≤ 20 dB HL at all frequencies—with the exception of one participant measuring 25 dB HL at 8000 Hz in both ears—and interaurally symmetrical (no interaural threshold difference > 15 dB HL at any frequency; Helfer and Freyman, 2008). Thresholds were generally higher and more variable among the older participants, but relatively well-preserved across the lower frequencies (mean thresholds ≤ 20 dB HL at ≤ 2000 Hz) with some mild hearing loss exhibited at some lower frequencies for some participants. Across the higher frequencies, thresholds for the older participants were characterized by an increasing slope that is common to aging (Hannula *et al.*, 2011; Hoffman *et al.*, 2012). A mixed analysis of variance (ANOVA) with the between-subjects factor age group (younger, older) and the within-subjects factor frequency range (low: 250–2000 Hz, high: 2000–8000 Hz) confirmed that thresholds were higher for the older group [$F(1, 42) = 119.99$, $p < 0.001$, $\eta_p^2 = 0.74$], and that group differences in hearing were larger for the higher frequency range [$F(1, 42) = 74.27$, $p < 0.001$, $\eta_p^2 = 0.64$]. Audiograms for the older participants can be described as “age-typical” since most thresholds did not exceed the range expected for 95% of the population based on age and gender (ISO, 2017). One notable exception was participant O7 (Fig. 1, individual audiograms for older participants) who exhibited more pronounced hearing loss across the lower frequencies. Three older participants showed some degree of asymmetrical hearing at one or more frequencies ≤ 2000 Hz (O1, O5, and O7). The inclusion of data from these potential outliers is discussed in greater detail in Sec. III.

C. Stimuli

Auditory stimuli consisted of low-pass filtered noise-vocoded speech in which single-syllable consonant-vowel-consonant (CVC) target words were presented with two-talker masking babble in the F-F, F-RF, and F-SUM configurations.

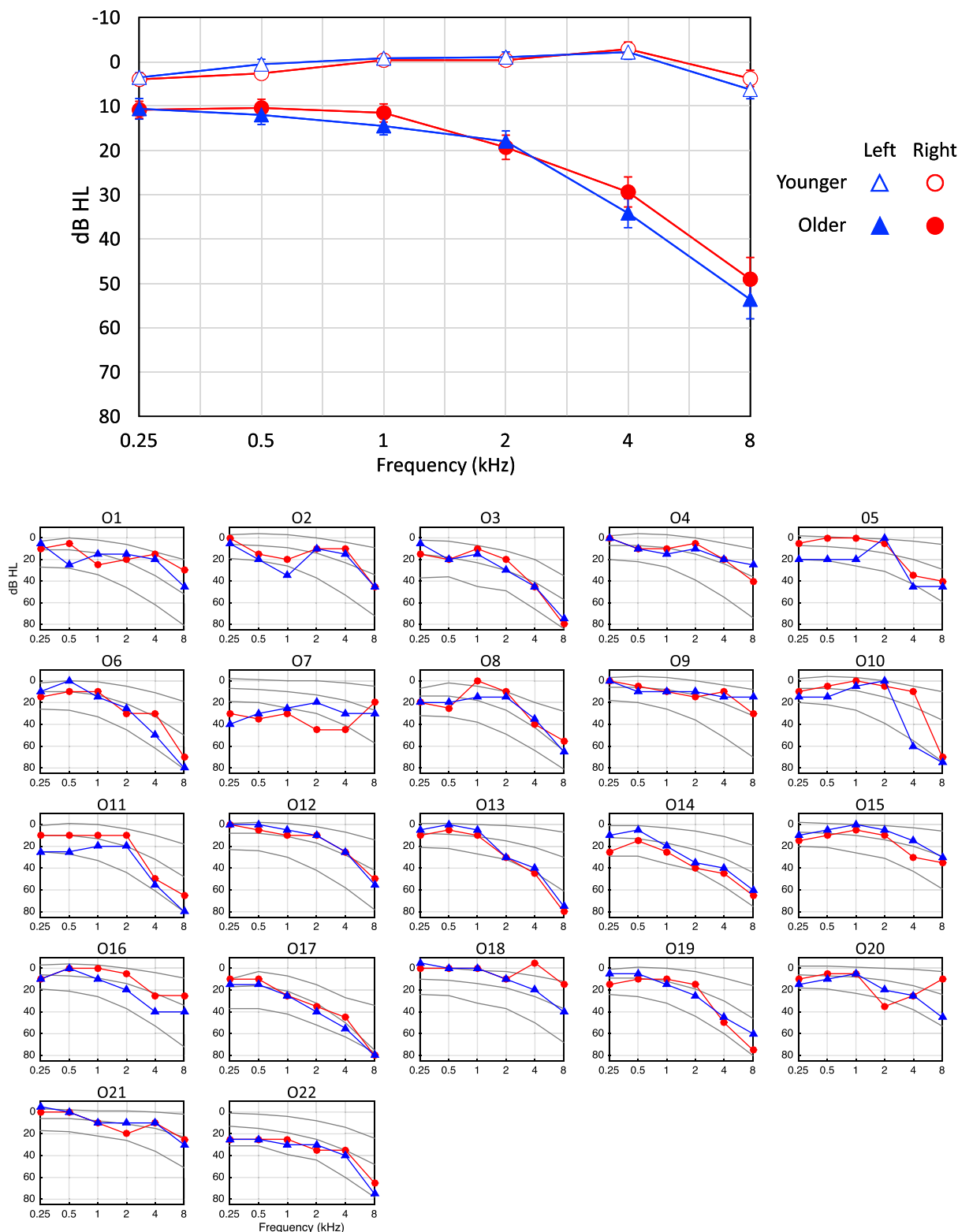


FIG. 1. (Color online) Top panel: Mean (± 1 standard error) pure-tone audiometric thresholds for both ears for the younger and older groups. typical of aging, the older group exhibited elevated thresholds compared to the younger group that were increasingly prominent across the higher frequencies. Bottom panel: audiograms of both ears for each older (O) participant. The three light gray lines show the median and lower and upper bounds of the estimated population distribution based on participant age and gender (5% of the population expected to fall below the lower bound and 5% above the upper bound). Hearing loss among the older participants was considered age-typical insofar as thresholds generally did not exceed the upper bound.

1. Target stimuli

Targets consisted of 70 words selected from a female talker's (average $F0 = 211.15$ Hz; [Boersma and Weenink, 2018](#)) recording of the Nederlandse Vereniging voor Audiologie

(NVA) list of common Dutch CVC words, widely used for speech audiometry in the Netherlands. Words with sharp acoustic onsets ([b], [d], [k], [p], [t]) were chosen to facilitate comparisons with previous ([Zobel et al., 2018](#)) and planned

electrophysiological studies. Likewise, in accordance with the relevant research (Freyman *et al.*, 2008; Morse-Fortier *et al.*, 2017; Zobel *et al.*, 2018), each target was noise vocoded with MATLAB (MathWorks Inc., 2015) using the procedure described in Qin and Oxenham (2003). First, a sixth-order Butterworth band-pass filter was used to divide the target into six contiguous bandwidths between 80 and 6000 Hz according to the Equivalent Rectangular Bandwidth scale designed to approximate the shape of human auditory filters (Glasberg and Moore, 1990). The envelope in each band was then extracted by low-pass filtering (second-order Butterworth) the half-wave-rectified band with a cutoff frequency set to the lower of either half the channel bandwidth or 300 Hz.² For synthesis bands, Gaussian white noise was bandpass filtered into the same six channels, and each channel of noise was modulated with its respective channel's extracted envelope. The resulting six channels were summed to create the vocoded version of the target.

2. Masker stimuli

Maskers consisted of two-talker female babble. The recordings were obtained from two separate corpora designed for measuring speech reception thresholds (Plomp and Mimpen, 1979; Versfeld *et al.*, 2000). The corpora, spoken by different female talkers (average F_0 s = 234.80 and 179.94 Hz, respectively; Boersma and Weenink, 2018) each consist of a series of simple, conversational Dutch sentences (130 and 507 sentences, respectively) describing everyday situations (e.g., “The ball flew over the fence”). For each corpus, all of the sentences were concatenated into a continuous stream that was edited such that no silence between words exceeded 100 ms in length. Each stream was then noise vocoded according to the procedure described above. Following vocoding, each stream was divided into 640 2.5-s one-talker segments. All one-talker segments were individually scaled to the same root-mean-square (RMS) amplitude. Then, each segment from one talker was summed with a randomly chosen segment from the other talker to create 640 2.5-s two-talker masker segments. The two-talker masker segments were individually scaled to the same RMS amplitude (masker RMS amplitude), which was held constant throughout the study.

3. Stimulus conditions

Fifty-six copies of each target were created and their RMS amplitudes were scaled in 1-dB steps from -40 to +15 dB relative to the two-talker masker RMS amplitude. Targets were saved as individual stereo files with the target placed in channel 1 and silence placed in channel 2.

Three versions of each masker were created as stereo wave files, consistent with the three spatial conditions to be tested. The F-F masker consisted of a single two-talker masker segment placed in channel 1 and silence placed in channel 2. The F-RF masker consisted of identical masker segments placed in both channels such that the onset of the segment in channel 2 preceded the onset of the segment in channel 1 by 4 ms. The F-SUM masker consisted of identical masker segments placed in both channels with synchronous onsets. Note that in all spatial conditions, SNR was

calculated as the RMS amplitude of the target in the front channel relative to the two-talker masker segment in the front channel.

4. Hearing-loss control

After assembling the targets and maskers, a 12th-order zero-phase Butterworth low-pass filter with a cutoff frequency of 2 kHz was applied to all stimuli (*filtfilt* function; MATLAB, MathWorks Inc., 2015). Figure 2 shows the long-term average spectra (LTAS) (Hummerson, 2017) of the low-pass filtered targets and maskers. Note that the low-pass filter was applied after the masker RMS amplitudes were equalized and target RMS amplitudes were scaled. Therefore, the actual SNRs varied slightly (masker and target SDs <1 dB RMS) around the SNR labels, consistent with the sensory response to unfiltered stimuli in an individual with profound high-frequency hearing loss.

D. Experimental setup

Figure 3 shows the configuration of the testing room. Participants were seated in a comfortable chair in the center of an electrically shielded 4.2 m × 2.5 m sound booth designed by Electro Medical Instruments to conform to ISO (2010) standards. Two shielded Yamaha HS8 loudspeakers were placed 1.4 m apart, with one loudspeaker at a distance of 1.4 m directly in front of the participant, and the other loudspeaker at a distance of 1.4 m and a horizontal angle of 60° to the participant's right. A computer screen was positioned just below the front loudspeaker to display text (e.g., instructions, fixation cross, response prompt) as white letters against a black background. Text was displayed at the top of the screen to keep the participant's head oriented with the axis of the front loudspeaker. Using E-PRIME (Psychology Software Tools, Inc., 2016) software, stimuli were presented as stereo WAV files with 16-bit resolution and 48 kHz sampling rate through a MOTU Ultralite-mk4 USB sound card. A Lavry DA10 digital-to-analog converter routed channel 1

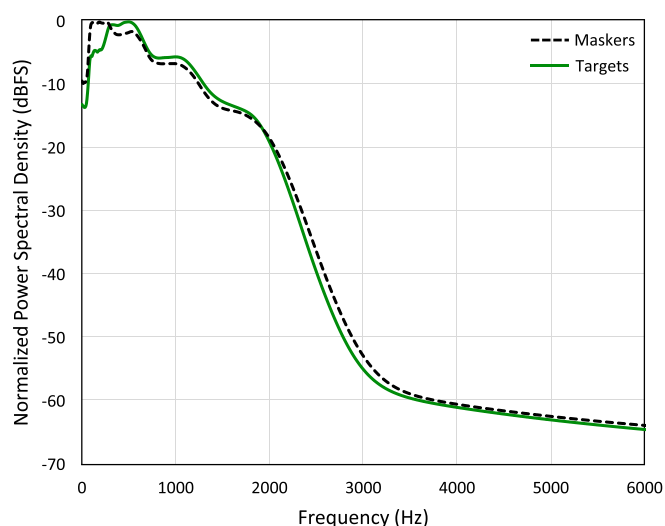


FIG. 2. (Color online) Long-term average spectra of the noise-vocoded maskers and targets. The sharp roll-off at 2 kHz reflects the application of the low-pass filter to control for hearing loss (12th-order zero-phase Butterworth).

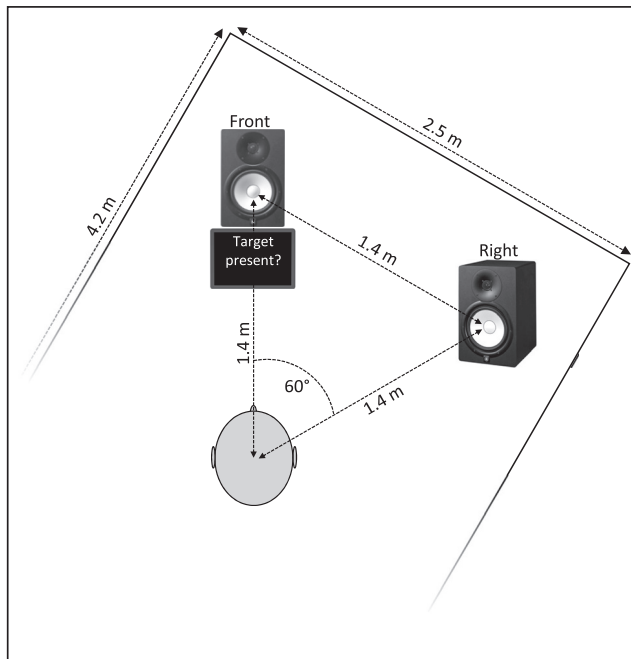


FIG. 3. Diagram of the experimental setup. Listeners were seated in the center of the room with the front and right loudspeakers facing them. A computer screen was placed just under the front loudspeaker to display text (e.g., response prompt). Targets were always presented from the front loudspeaker while masker presentation differed across spatial conditions. In the F-F condition, maskers were presented from the front loudspeaker only, and no stimulus was presented from the right loudspeaker. In the F-RF condition, identical maskers were presented from both loudspeakers with the onset of the right masker preceding the onset of the front masker by 4-ms. In the F-SUM condition, identical maskers were presented from both loudspeakers with synchronous onsets.

of the stereo files to the front loudspeaker and channel 2 to the right loudspeaker. Prior to beginning the study, the gain for each loudspeaker was individually adjusted to equate their sound levels; the average level of a stream of masker segments presented from a loudspeaker was 70 dBA at the position of the listener's head when measured on-axis with the loudspeaker using a Svantek 979 sound level meter.

E. Procedure

Following pure-tone audiometry and the Mini-Mental State Examination (older group only), participants began the target-detection task. On each trial, a masker was presented, followed 500–1500 ms later (interval randomly chosen on each trial in ms resolution) by either the presentation of a target or no presentation of a target. A fixation cross appeared on the screen 500 ms before the masker onset, and remained for 500 ms after the masker offset, followed by a response prompt which asked participants to press a button indicating whether a target had been present on the trial (“yes” response), or not (“no” response). No feedback was provided during the experimental trials.

Prior to beginning the task, participants received instructions. They were told that on each trial they would be deciding whether or not a “target voice” was present among “other voices.” They were told that when the target voice was present, it would always come from the front loudspeaker and would

only say a single word. They were also explicitly told that their task was not to understand what the target voice was saying, but to simply judge whether or not the target voice was present on each trial. Participants were then presented with examples of targets in isolation and asked to confirm that they could hear each target from the front location. Next, examples of F-F, F-RF, and F-SUM maskers were presented in isolation, and participants were asked to confirm that they heard each masker from its respective location (i.e., front, right, and between front and right). They were then presented with clear examples of target-present trials (+10–12 dB SNRs) in the three spatial conditions and were asked to confirm that they could detect the targets. Participants then completed 15 practice trials consisting of three target-present (+10–12 dB SNRs) and two target-absent trials in each spatial condition while the experimenter watched to confirm that responses were consistent with understanding the task. Participants were told that on the real trials, it would not always be so clear as to whether or not the target was present, and that they should use their best judgment. They were also instructed to remain oriented toward the front loudspeaker with their eyes on the fixation cross while listening. Participants were otherwise left free to adopt any listening strategy for detecting the target voice, which may have included listening for fluctuations in amplitude or disruptions in the patterns of the sounds.

Following these instructions, participants completed 630 experimental trials comprising 30 trials at each of seven SNRs within each of the three spatial conditions. The seven SNRs, which included an SNR designated for target-absent trials (SNR_{NULL}), were chosen for each spatial condition to cover the relevant range of the psychometric functions that were predicted to be obtained based on pilot data (F-F: +15, +5, 0, –5, –10, –15 dB, and SNR_{NULL} ; F-RF and F-SUM: 0, –10, –15, –20, –25, –30 dB, and SNR_{NULL}). The trials were divided into six blocks of 105 trials (5 trials \times 7 SNRs \times 3 spatial conditions) presented in random order. The masker presented on each trial was randomly selected from the 640 maskers available in each spatial condition such that a two-talker masker segment could not be presented more than once within a given spatial condition in the same block. Likewise, the target word presented on each target-present trial was randomly chosen from the 70 available target words such that a target word could not be presented more than once within a given spatial condition in the same block. The experiment took approximately two hours to complete.

F. Data analysis

1. Detection thresholds

Detection rates in a yes/no task reflect independent contributions of accuracy and response bias, according to the firmly-established Signal Detection Theory (Green and Swets, 1966; Macmillan and Creelman, 2005). Therefore, a measure of each participant's detection threshold (accuracy), independent of their response bias, was estimated from the data in each spatial condition. Spatial release from masking could then be calculated as the change in detection threshold between the spatially co-located and spatially separated conditions to determine whether age-group differences were

observed (research question 1) and, if so, whether age independently of hearing loss accounted for any of the variability (research question 2).

To estimate detection thresholds, the psychometric function developed by Lesmes *et al.* (2015) was fit to each participant's detection rates in each spatial condition. This model was implemented early in the process of designing the present study, when a Bayesian adaptive yes/no task was initially considered. Although the method of constant stimuli (i.e., collecting detection rates across a range of stimulus intensities) was ultimately chosen for the present study, model comparisons (Lesmes *et al.*, 2015, appendixes) and analysis of pilot data indicated that the model—deeply rooted in the theoretical and empirical applications of Signal Detection Theory—would provide a good fit. At the core of the model is the d' function adapted from Lesmes *et al.* (2015) such that d' at any SNR is given by

$$d'(SNR) = \frac{\beta(SNR/\tau)^\gamma}{\sqrt{(\beta^2 - 1) + (SNR/\tau)^{2\gamma}}}, \quad (1)$$

with β determining the d' value at which the function asymptotes, τ representing the detection threshold, and γ determining the slope of the function. Under this formulation, the detection threshold (τ) is defined as the SNR at which d' will be equal to 1. This is equivalent to a score of 76% correct in an analogous two-alternative forced choice (2AFC) task (Stanislaw and Todorov, 1999). SNRs are entered in linear units of amplitude, but the abscissa of the d' function is in relative units (i.e., units of threshold) that are convertible to decibels (Klein, 2001; Lesmes *et al.*, 2015).

The psychometric function, adapted from Lesmes *et al.* (2015), uses the d' function to obtain the detection rate (Ψ_{yes} ; proportion of “yes” responses) at any given SNR with

$$\Psi_{\text{yes}}(SNR) = 1 - G(\lambda - d'(SNR)), \quad (2)$$

where $G(x)$ is the standard normal cumulative distribution function and λ is the measure of response bias. Additionally, a lapse rate (ε), accounting for the proportion of trials on which stimulus-independent behavioral lapses (blinks, distracted attention, response errors, etc.) are expected to occur, is incorporated into the model to improve the accuracy of parameter estimation (Lesmes *et al.*, 2015; Wichmann and Hill, 2001). The model assumes an equal distribution of “yes” and “no” responses on lapse trials. Thus, the final form of the psychometric function used for the present study, adapted from Lesmes *et al.* (2015), is given by³

$$\Psi'_{\text{yes}}(SNR) = \frac{\varepsilon}{2} + (1 - \varepsilon)\Psi_{\text{yes}}(SNR). \quad (3)$$

For each participant, the psychometric function (Ψ'_{yes}) was fit to the detection rates obtained at the seven SNRs in each spatial condition, allowing the detection threshold (τ), slope (γ), and response bias (λ) parameters to vary freely. The d' function's asymptote (β) was fixed at 5, and the lapse rate (ε) was fixed at 0.01 (Lesmes *et al.*, 2015; Wichmann and Hill, 2001). Consistent with Lesmes *et al.* (2015), the detection

rates predicted by the psychometric function ($P_{\text{predicted}}$) were fit to the participant's observed detection rates (P_{observed}) at the seven SNRs with the set of parameter values that minimized the Pearson's χ^2 statistic given by

$$\chi^2 = \sum_{SNR} \frac{(P_{\text{observed}} - P_{\text{predicted}})^2}{[P_{\text{predicted}} \times (1 - P_{\text{predicted}})]/n}, \quad (4)$$

with n set to the number of trials (i.e., 30 trials) at each SNR (Lesmes *et al.*, 2015; Wichmann and Hill, 2001). To avoid local minima, a two-step routine was carried out in which initial minimization with a broad grid search was used to identify the best set of parameter values to be entered as starting points for subsequent minimization with the *fminsearch* function in MATLAB (MathWorks Inc., 2015).

2. Statistical analyses

The planned statistical analyses were designed to answer the two research questions stated above, and were conducted in IBM SPSS STATISTICS FOR MACINTOSH (IBM, 2015). To assess differences in spatial release from informational masking between the two age groups (question 1), analyses of the behavioral data were first conducted to describe any observed patterns of group-based differences in detection rates across the SNRs in the three spatial conditions. Detection rates were entered into mixed ANOVAs with age group as the between-subjects factor, and spatial condition and/or SNR as the within-subjects factors. To confirm that any observed behavioral effects were driven by group-based differences in detection accuracy, independent of response bias, the detection thresholds estimated for each participant were entered into a mixed ANOVA with age group as the between-subjects factor and spatial condition as the within-subjects factor. Follow-up analyses on the detection-rate and threshold data were conducted when important between- and within-subject main effects and interactions were indicated by the omnibus ANOVAs. While the uncorrected degrees of freedom are reported, the Greenhouse-Geisser correction was applied to the p-values when violations of sphericity were indicated by Mauchly's test. The p-values were also corrected in follow-up independent-sample t-tests when heterogeneity of variance was indicated by Levene's test.

To assess the extent to which age, independent of age-typical hearing loss, predicted the amount of spatial release from informational masking (question 2), multiple linear regression analysis was performed. Spatial release (spatially co-located detection threshold minus spatially separated detection threshold) was entered as the outcome variable, and age and hearing loss were entered as independent predictors. The measure of hearing loss was chosen *a priori* to be the pure-tone average (PTA) of both ears across the four stimulus frequencies shown to be most relevant in the LTAS (Fig. 2): 250, 500, 1000, and 2000 Hz. Interpretation of standardized effect sizes is limited by the fact that only younger and older adults were sampled; therefore, regression coefficients are reported in unstandardized units (Preacher *et al.*, 2005).

Analyses of slope and response bias estimates were not informative with regard to age-related differences in spatial

release from informational masking, and are not presented here for the sake of simplifying the reported results.

III. RESULTS

A. Detection rates

Figure 4 presents the mean detection rates (i.e., proportion “yes” responses) of the younger and older groups in the three spatial conditions. The classic S-shaped pattern of the data confirms that the range of SNRs were well-chosen to cover the relevant extent of the psychophysical responses within each spatial condition. Since the SNRs tested in the F-F condition differed from those tested in the F-RF and F-SUM conditions, separate detection-rate analyses were

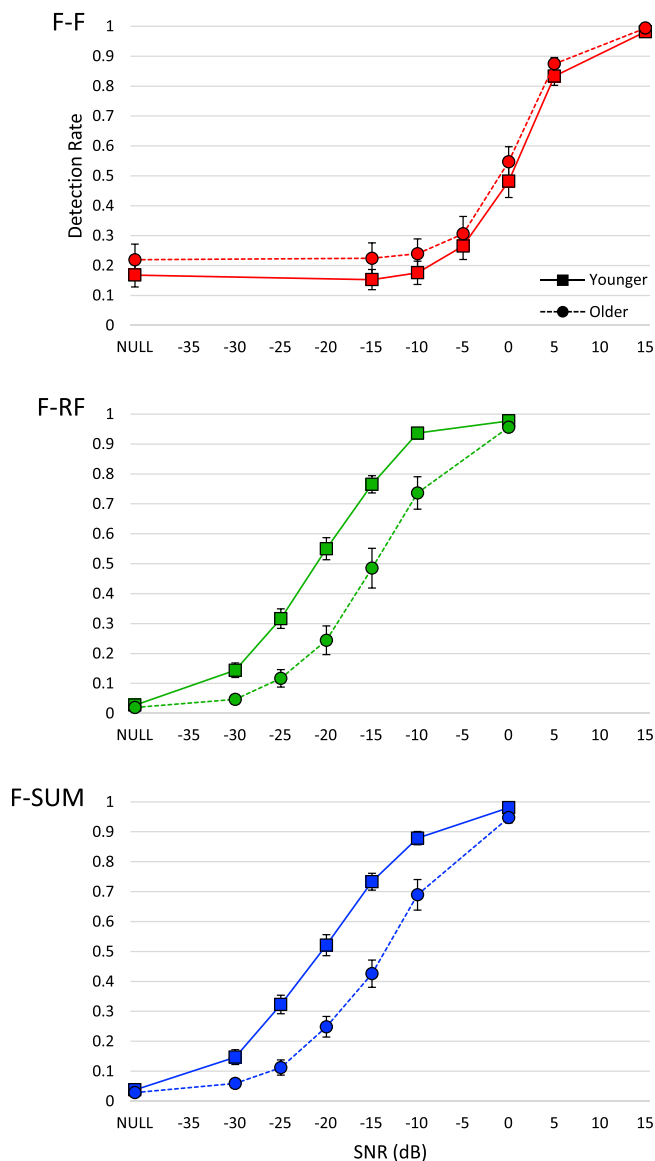


FIG. 4. (Color online) Mean (± 1 standard error) detection rates (proportion of trials eliciting a “yes” response) at the seven SNRs in the three spatial conditions for the younger and older groups. No age-group differences were found in the spatially co-located condition (F-F). In the spatially separated conditions (F-RF and F-SUM), detection rates were similar between the age groups at the extreme SNRs (SNR_{NULL} and SNR_0), but were otherwise reduced for the older group compared to the younger group, consistent with a reduction in spatial release from informational masking.

conducted in the spatially co-located and separated conditions. In the F-F condition, the detection rates and shape of the psychophysical data were remarkably similar for the younger and older groups. Detection rates analyzed with a 2 age group (younger, older) \times 7 SNRs (+15, +5, 0, -5, -10, -15 dB, and SNR_{NULL}) mixed ANOVA did not indicate any difference between the responses of the younger and older groups (main effect and interaction: $p \geq 0.32$), nor did independent-sample t-tests conducted at each SNR ($p \geq 0.24$). In comparison to the F-F condition, data in the F-RF and F-SUM conditions showed a strikingly different pattern. Data were analyzed with a 2 age group \times 2 spatial condition (F-RF, F-SUM) \times 7 SNR (0, -10, -15, -20, -25, -30 dB, SNR_{NULL}) mixed ANOVA. A significant main effect of age group [$F(1, 42) = 26.57, p < 0.001, \eta_p^2 \geq 0.39$] and an interaction between age group and SNR [$F(6, 252) = 16.11, p < 0.001, \eta_p^2 \geq 0.28$] were driven by the fact that detection rates were similar between the age groups at the extreme SNRs ($\text{SNR}_0, \text{SNR}_{\text{NULL}}$: t-test $p \geq 0.09$), but substantially lower for the older group at the SNRs in between [$t(42) \geq 3.24, p \leq 0.002, d \geq 0.98$], suggesting poorer accuracy. Furthermore, there was no indication that these age-group effects differed between the F-RF and F-SUM conditions (age group \times spatial condition \times SNR interaction: $p = 0.73$) or that detection rates, when analyzed separately within each group, differed between the F-RF and F-SUM conditions (spatial condition main effect and interaction with SNR within younger and older groups: $p \geq 0.10$).

B. Threshold estimates

To further examine the differences in performance between age groups that were indicated by the global characteristics of the psychophysical data, analyses were performed on the thresholds estimated by the psychometric functions fit to each participant’s detection rates in the three spatial conditions. Figures 5 and 6 show that the psychometric model captured the detection rates well in each spatial condition, given a critical $\chi^2(3)$ of 7.81 at $\alpha = 0.05$ [mean $\chi^2_{F-F}(3) = 3.50, SD = 2.87$; mean $\chi^2_{F-RF}(3) = 3.75, SD = 2.80$; mean $\chi^2_{F-SUM}(3) = 3.43, SD = 2.49$]. No difference in goodness of model fit was found between age groups either within or across spatial conditions ($p \geq 0.11$). One older participant (O6 in Fig. 6) was a notable outlier in the F-F condition, with a threshold estimate (13.21 dB SNR) that was 3.93 standard deviations above the mean of the older group, whose thresholds otherwise ranged from -4.54 to 1.93 dB SNR. Potential outliers are addressed in greater detail below.

Figure 7 presents the mean thresholds calculated at $d' = 1$, when target and masker were spatially co-located (F-F), and spatially separated (F-RF, F-SUM). Thresholds were nearly identical for the younger and older groups in the F-F condition, and were markedly reduced for both age groups in the spatially separated conditions. F-RF and F-SUM thresholds were similar within each age group but elevated for the older group compared to the younger group. Threshold measurements for each participant were entered into a 2 age group (younger, older) \times 3 spatial condition (F-F, F-RF, F-SUM) mixed ANOVA. Significant main effects of age group

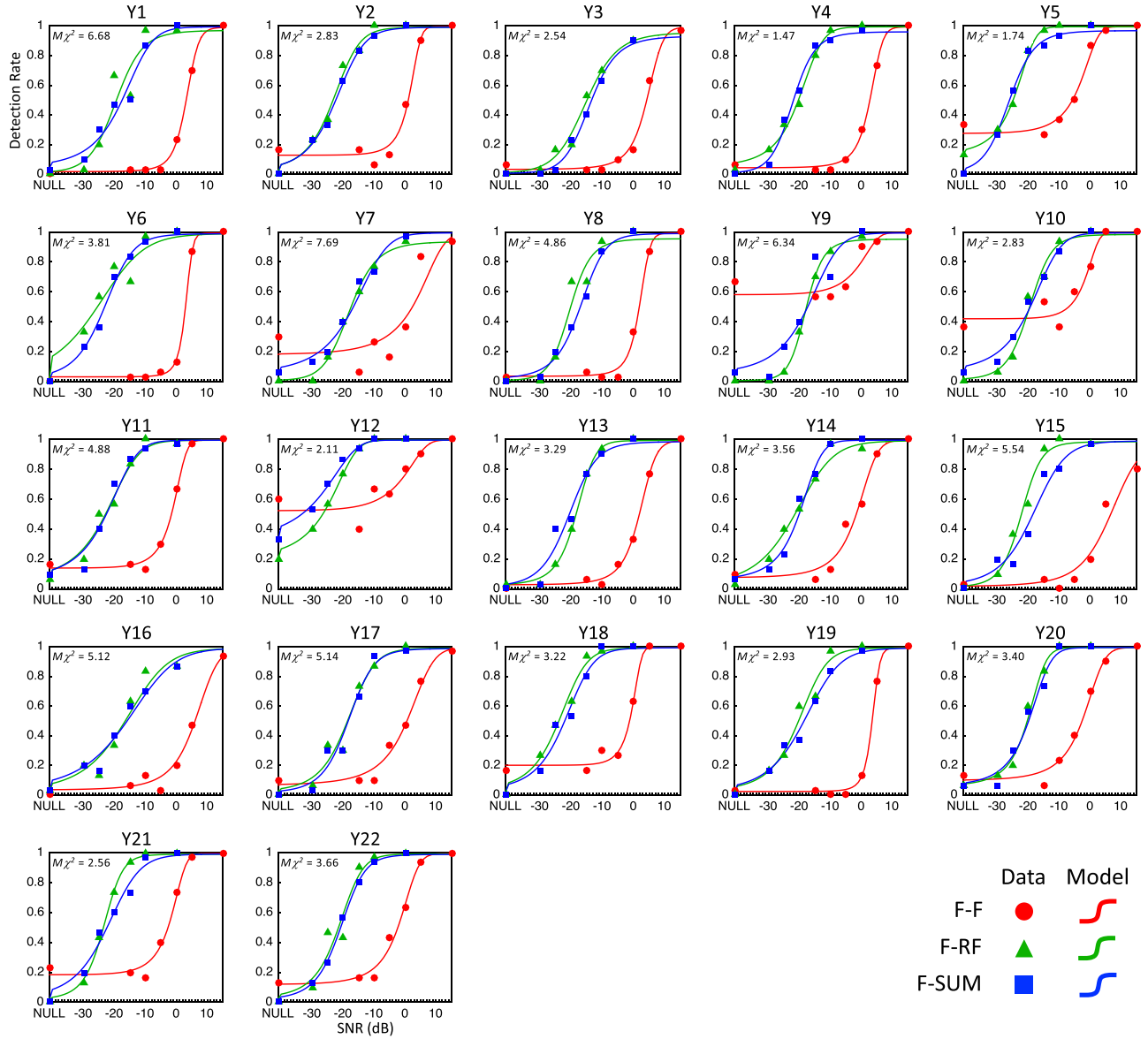


FIG. 5. (Color online) Psychometric functions (curves) fit to the data (points) of the younger participants. The mean of the Pearson's χ^2 fit statistics ($M\chi^2$) for the F-F, F-RF, and F-SUM functions is included in each plot [critical $\chi^2(3)=7.81$ at $\alpha=0.05$].

[$F(1, 42)=24.76$, $p<0.001$, $\eta_p^2=0.37$] and spatial condition [$F(2, 84)=399.28$, $p<0.001$, $\eta_p^2=0.91$], and an age group \times spatial condition interaction [$F(2, 84)=7.61$, $p=0.001$, $\eta_p^2=0.15$] supported the observation that substantial spatial release from masking was exhibited by both groups but was reduced for the older group compared to the younger group. Indeed, follow-up independent-sample t-tests comparing the age groups in each spatial condition did not find that thresholds differed in the F-F condition ($p=0.41$), while thresholds were higher for the older group compared to the younger group in both the F-RF [$t(42)=4.41$, $p<0.001$, $d=1.33$] and F-SUM [$t(42)=3.90$, $p<0.001$, $d=1.18$] conditions. To investigate whether these age-group effects differed between the spatially separated conditions, thresholds were entered into a 2 age group \times 2 spatial condition (F-RF, F-SUM) mixed ANOVA. No main effect of spatial condition and no interaction between age group and spatial condition was found ($p\geq 0.37$). In addition, repeated-

measures ANOVAs performed separately within each age group showed clear spatial release from masking in the F-RF [younger: $F(1, 21)=528.25$, $p<0.001$, $\eta_p^2=0.96$; older: $F(1, 21)=212.51$, $p<0.001$, $\eta_p^2=0.91$] and F-SUM [younger: $F(1, 21)=653.89$, $p<0.001$, $\eta_p^2=0.97$; older: $F(1, 21)=138.05$, $p<0.001$, $\eta_p^2=0.87$] conditions, while there was no indication that thresholds differed between the spatially separated conditions for either the younger group ($p=0.56$) or the older group ($p=0.50$). Taken together, these results show that masking release in the F-RF and F-SUM conditions was similar for participants within each group, but substantially reduced in the older group compared to the younger group.

C. Independent contributions of age and hearing loss

To investigate the independent contributions of aging and hearing loss in predicting the reduced spatial release

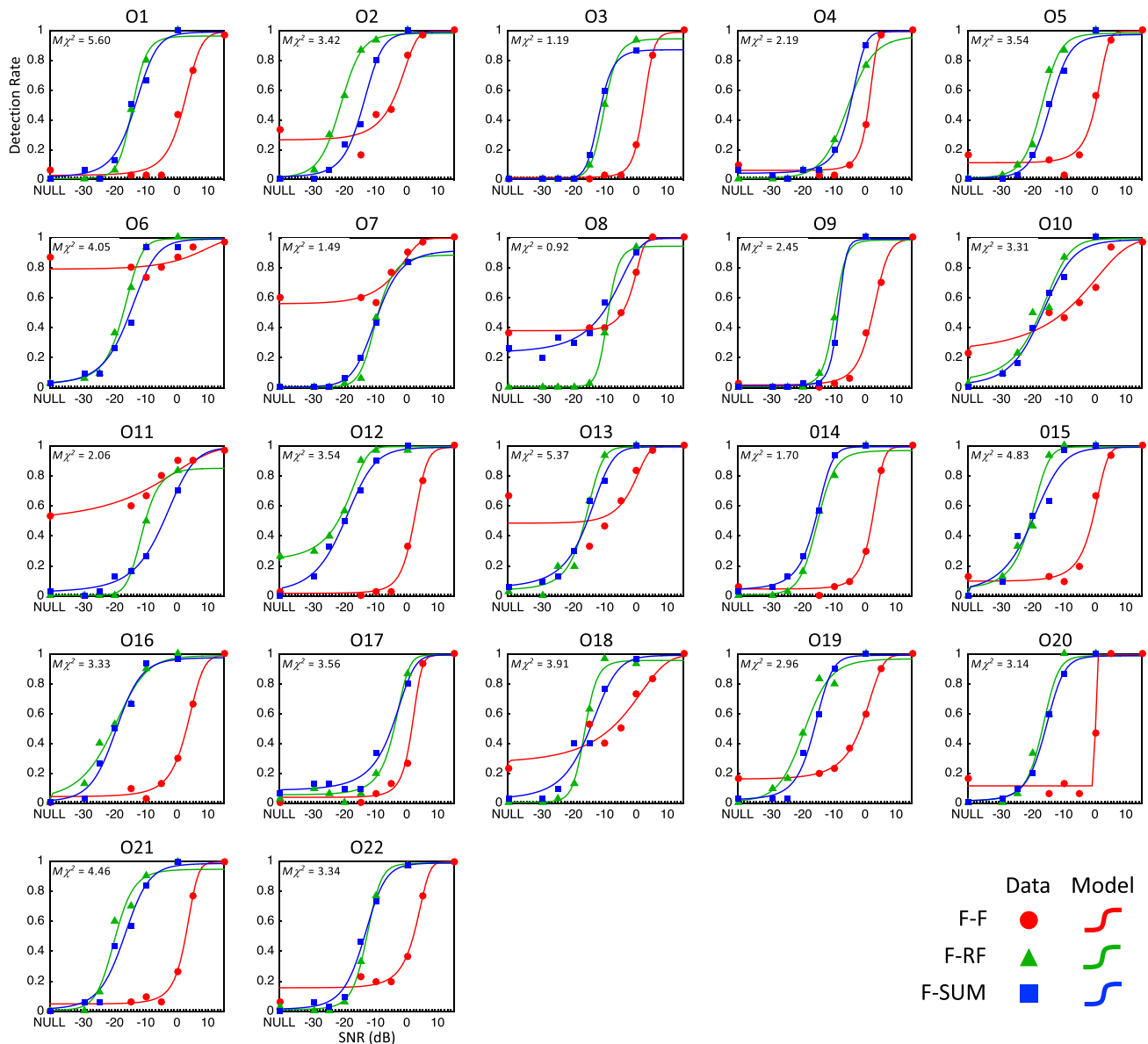


FIG. 6. (Color online) Psychometric functions (curves) fit to the data (points) of the older participants. The mean of the Pearson's χ^2 fit statistics ($M\chi^2$) for the F-F, F-RF, and F-SUM functions is included in each plot [critical $\chi^2(3) = 7.81$ at $\alpha = 0.05$].

from masking observed in the older group, multiple linear regression analysis was performed. The initial analysis entered age as a dichotomous predictor (young, old), and hearing loss as a continuous predictor. To increase power and reduce the number of analyses, a single outcome measure, spatial release, was calculated by subtracting for each participant the average of their F-RF and F-SUM thresholds from their baseline F-F threshold. This approach was supported by the fact that there was (1) no indication that thresholds differed between younger and older adults in the F-F condition, (2) no indication that the difference observed between age groups was different between the F-RF and F-SUM conditions, and (3) no indication that performance within each age group differed in the F-RF and F-SUM conditions. A positive correlation was found between age and hearing loss [$r_{pb}(42) = 0.74$, $p < 0.001$], while negative correlations were found between age and spatial release [$r_{pb}(42) = -0.52$, $p < 0.001$], and hearing loss and spatial

release [$r(42) = -0.46$, $p = 0.002$]. Results of the regression analysis showed a decrease in spatial release for the older group compared to the younger group independent of hearing loss ($b = -5.71$, $SE = 2.83$, $p = 0.05$, $sr = -0.27$), while hearing loss was not shown to predict spatial release independent of age group ($b = -0.15$, $SE = 0.17$, $p = 0.39$, $sr = -0.12$).

Given that the age ranges were rather broad within the younger and older groups (range = 16 and 20 years, respectively), exploratory analyses were conducted to determine the extent to which age-related declines in spatial release could be detected within each group. Among the older participants, a significant positive correlation between age and hearing loss was found [$r(20) = 0.49$, $p = 0.02$]. Age was significantly negatively correlated with spatial release [$r(20) = -0.43$, $p = 0.05$], but hearing loss was not [$r(20) = -0.16$, $p = 0.49$]. Furthermore, when spatial release was regressed on age and hearing loss, age was shown to be

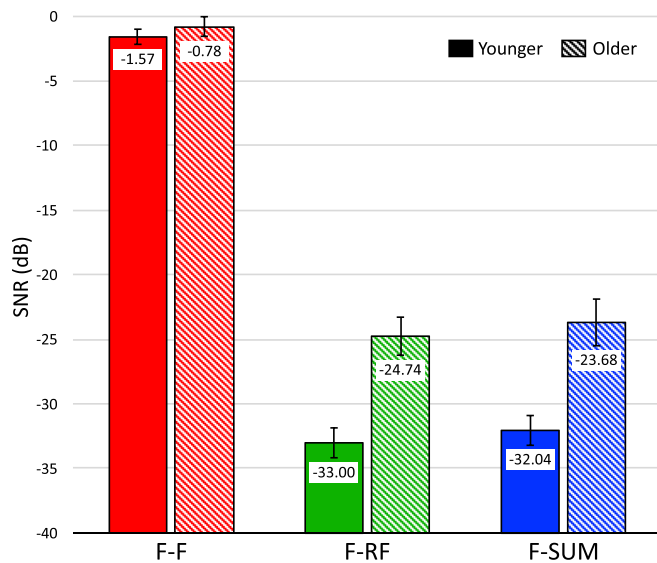


FIG. 7. (Color online) Mean (± 1 standard error) thresholds ($d' = 1$) for the younger and older groups in the spatially co-located (F-F) and spatially separated (F-RF, F-SUM) conditions. Thresholds were not found to differ between the age groups in the F-F condition. Both groups exhibited spatial release from masking (co-located threshold minus spatially separated threshold), but the masking release was reduced for the older group compared to the younger group.

marginally associated with reduced spatial release ($b = -0.59$, $SE = 0.30$, $p = 0.07$, $sr = -0.41$), and significantly associated with reduced spatial release when the outlier in the F-F condition (O6 in Fig. 6) was removed from analysis ($b = -0.65$, $SE = 0.26$, $p = 0.02$, $sr = -0.50$). In contrast, hearing loss failed to significantly predict spatial release among the older participants in either model when controlling for age (with/without outlier: $b = 0.07/0.08$, $SE = 0.24/0.21$, $p = 0.77/0.70$, $sr = 0.06/0.08$). Interestingly, a similar pattern of results was indicated within the younger group. Age and hearing loss were not found to be correlated [$r(20) < 0.001$, $p > 0.99$], but a marginal negative correlation between age and spatial release was found [$r(20) = -0.39$, $p = 0.08$] while no such correlation was indicated for hearing loss [$r(20) = -0.06$, $p = 0.80$]. Furthermore, regression analysis showed that age marginally predicted a

decrease in spatial release independent of hearing loss ($b = -0.42$, $SE = 0.23$, $p = 0.08$, $sr = -0.39$), while there was no indication that hearing loss predicted spatial release independent of age ($b = -0.08$, $SE = 0.31$, $p = 0.79$, $sr = -0.06$).

The fact that regression analyses conducted within each group were consistent with analyses conducted across groups when age was dichotomously collapsed, suggested a robust age-related decline in spatial release best described as a continuous measure. Across all participants, a continuous measure of age was positively correlated with hearing loss [$r(42) = 0.77$, $p < 0.001$] and negatively correlated with spatial release [$r(42) = -0.59$, $p < 0.001$]. As shown in Fig. 8, regressing spatial release on continuous measures of age and hearing loss showed an age-related decline in spatial release independent of hearing loss ($b = -0.18$, $SE = 0.06$, $p = 0.007$, $sr = -0.36$). In contrast, there was no indication that hearing loss predicted spatial release independent of age ($b = -0.02$, $SE = 0.17$, $p = 0.93$, $sr = -0.01$). To confirm that the latter result was not solely dependent upon the lower frequency range of the hearing loss measure, the same regression analysis was conducted with hearing loss calculated across the higher frequencies where age-group differences were more pronounced (PTA across 2000, 4000, 8000 Hz) and hearing loss showed a stronger positive correlation with age [$r(42) = 0.90$, $p < 0.001$] and negative correlation with spatial release [$r(42) = -0.50$, $p = 0.001$]. Regression analysis confirmed an age-related decline in spatial release independent of hearing loss ($b = -0.24$, $SE = 0.09$, $p = 0.01$, $sr = -0.32$), while there was no indication that hearing loss predicted spatial release independent of age ($b = 0.07$, $SE = 0.11$, $p = 0.52$, $sr = 0.08$).⁴

D. Potential outliers

1. Audiometric outliers

As previously discussed, Participant O7 (Fig. 1) exhibited hearing loss across the lower frequencies that was somewhat more pronounced and less typical compared to the other older adults. However, data from O7 was included in analysis after determining that O7 actually exhibited greater

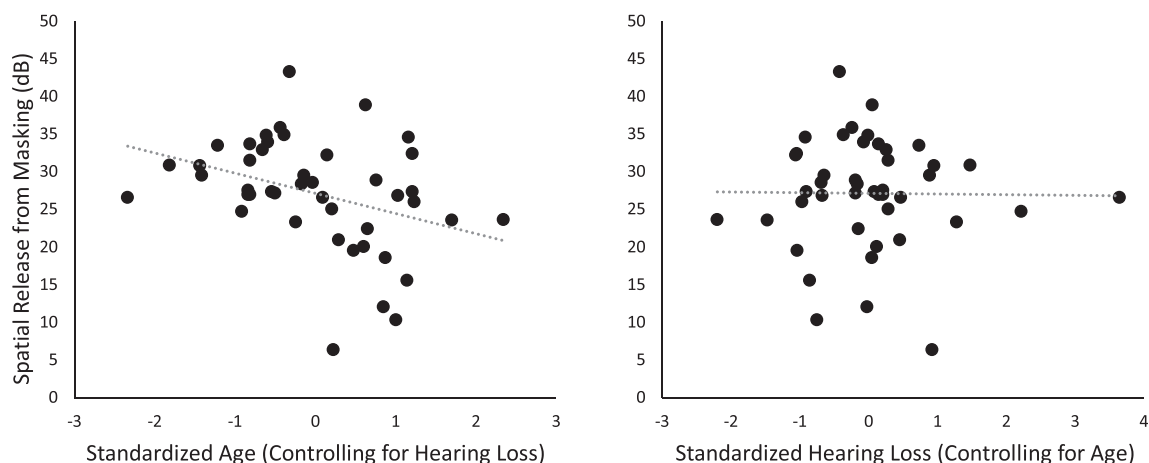


FIG. 8. Independent contributions of age and hearing loss in predicting spatial release from masking. Age predicted a decline in spatial release from masking, independent of hearing loss (left panel), while there was no indication of an independent relationship between hearing loss and spatial release from masking (right panel). X-axes are the standardized residuals of age regressed on hearing loss (left panel) and hearing loss regressed on age (right panel).

spatial release from masking than the mean of the older group ($+0.42$ *SD* spatial release), working against the reported effects, and that excluding them from analysis did not substantively change the key results presented in Figs. 4, 7, and 8. Participant O7, along with O1 and O5 (Fig. 1) also exhibited some degree of interaural asymmetry (>15 dB HL) among frequencies ≤ 2000 Hz. Again, data from these participants were included in analysis after finding that they, as a group, exhibited slightly greater mean spatial release from masking compared to the older group ($+0.14$ *SD* spatial release), and that excluding them from analysis did not substantively change the key results. The definition of asymmetrical hearing, however, varies across the relevant literature (e.g., Gallun *et al.*, 2013; Helfer and Freyman, 2008), and is not agreed upon in the clinical literature (Saliba *et al.*, 2011). Therefore, to more broadly rule out the influence of any degree of asymmetrical hearing across participants, separate analyses not reported here were conducted in which a measure of asymmetry (sum of the interaural variances in thresholds calculated at 250–8000 Hz) was included as an independent variable in the regression models described above. Degree of asymmetrical hearing was not shown to significantly predict spatial release nor substantively influence the reported results.

2. Threshold outlier

As previously discussed, the threshold estimate for participant O6 far exceeded those of the older group in the F-F condition, while O6's threshold estimates in the spatially separated conditions did not (Fig. 6). Thus, a conservative approach was taken in deciding to include this participant in analysis, because their large release from masking only served to work against the reported effects. Excluding O6 in a separate analysis was not shown to substantively change the key results.

IV. DISCUSSION

The present study sheds light on age-related declines in spatial release from informational masking that may contribute to speech-processing difficulties under challenging listening conditions. The study was designed to address two fundamental questions: (1) Does spatial release from informational masking decline with age and, if so (2) does age predict this decline independently of age-typical hearing loss? The results provide support in answering “yes” to both questions. Although both age groups exhibited spatial release from informational masking, considerable reductions in masking release were observed among the older participants. These reductions were clear enough to be evident in the raw detection-rate data (Fig. 4) and further, were precisely described in the threshold data obtained from psychometric modeling (Figs. 5, 6, and 7). Additional regression analyses provided evidence that the observed age-related declines in spatial release from informational masking were independent of age-typical hearing loss (Fig. 8).

The clarity of the results obtained in the present study was a main goal of the experimental design. The use of the virtual separation paradigm was intended to isolate spatial

release to only the informational portion of the masking (i.e., portion of the masking related to target/masker confusion). The use of a detection task with low-pass filtered noise-vocoded stimuli was intended to reduce task demands and non-spatial differences between targets and maskers, to better control for age-typical hearing loss, to match performance between the age groups when target and masker were spatially co-located, and to increase the size of the measured masking release to better observe age-group differences in the basic mechanisms underlying spatial release from informational masking. These objectives were born out in the experimental results. The simplicity of the task also likely contributed to obtaining clear and consistent detection-rate data conducive to psychometric modeling (Figs. 5 and 6). The results show that a measure of accuracy could be obtained from the yes/no task. The large masking release observed was consistent with large effects reported in prior research using 4AFC target-detection tasks with noise-vocoded speech (Freyman *et al.*, 2008; Morse-Fortier *et al.*, 2017), supporting the validity of the current method and future use of yes/no paradigms, especially in light of their inherent advantages (Kaernbach, 1990; Klein, 2001). Moreover, the model of Lesmes *et al.* (2015) proved to be a good fit for the data based on both the statistical evidence, and on the fact that the threshold estimates were consistent with the patterns explicit in the raw detection-rate data.

Unlike much of the prior research that used natural speech identification tasks, the present study found nearly identical performance in the F-F condition for younger and older participants. Matched performance in the F-F condition suggests that the older participants did not experience greater amounts of informational masking than the younger participants when the target and masker were spatially co-located. These findings are consistent with Helfer and Freyman (2008), who manipulated the confusability of targets and maskers and concluded that older listeners do not exhibit increased susceptibility to informational masking compared to younger listeners. Susceptibility to informational masking was not directly tested in the present study; however, thresholds in the F-F condition were consistently close to 0 dB SNR, a region at which informational masking has been posited to reach a maximum in younger adults (Arbogast *et al.*, 2005; Freyman *et al.*, 2008). Therefore, it can be argued from the present study that the ceiling for informational masking does not appear to increase with age. Instead, there appears to be an age-related decrease in the amount of masking that is released when target and masker are perceived to be spatially separated.

Matched performance in the F-F condition allowed for an assessment of group-based differences in masking release across spatial conditions that avoided the assumptions and potential confounds of transforming data to equate performance (e.g., Helfer and Freyman, 2008; Li *et al.*, 2004). It is important to note that target-detection accuracy for both age groups was dramatically improved in the spatially separated conditions, suggesting that older adults continue to maintain heavy reliance upon spatial release from informational masking under challenging listening conditions. However, in sharp contrast to the matched performance observed in the

F-F condition, target-detection accuracy was markedly reduced for the older participants compared to the younger participants in the F-RF and F-SUM conditions, as evidenced by lower detection rates on target-present trials despite nearly identical inter-group false alarm rates (i.e., “yes” responses on target-absent trials), and by elevated threshold estimates. Crucially, the reduced accuracy for the older group was similar regardless of whether the target and masker were spatially separated by the precedence effect or summing localization, pointing to an age-related decline in the ability to benefit from the perception of spatial separation more generally, rather than a decline specific to the precedence effect or summing localization. On average, spatial release from informational masking was reduced in the older participants by 7.5 dB compared to the younger participants. Further research is required to determine how such a reduction in informational masking release under the present conditions may relate to speech-processing difficulties within the complex, multi-talker environments of everyday life.

In addition to demonstrating an age-related reduction in spatial release from informational masking, the present study provided evidence that this reduction was related to aging itself, independent of age-typical hearing loss (Fig. 8). This finding is consistent with Gallun *et al.* (2013) and Srinivasan *et al.* (2016) insofar as age was shown to significantly predict spatial release from masking, controlling for hearing loss. However, Gallun *et al.* (2013) found that age and hearing loss both independently predicted spatial release in some experiments, and Srinivasan *et al.* (2016) found that hearing loss was a dominant predictor of spatial release for large spatial separations. A large spatial separation was used in the present study, but only age independently predicted masking release; there was no indication of an age-independent relationship between hearing loss and masking release. One important difference is that Gallun *et al.* (2013) and Srinivasan *et al.* (2016) used physical spatial separation, which may have allowed energetic masking release modulated by hearing loss to influence results, while the use of virtual separation in the present study may have been better able to isolate effects related to informational masking release. However, another crucial point of consideration is offered by Srinivasan *et al.* (2016) who suggest that their inclusion of a hearing-impaired older group may have allowed effects of greater hearing loss to be revealed. Indeed, when Srinivasan *et al.* (2016) removed the hearing-impaired older group from analysis, and compared younger and older adults with hearing thresholds similar to the participants in the present study, only age was found to predict declines in spatial release from masking. Future research using the present paradigm will need to test hearing-impaired older adults to determine the extent to which greater degrees of hearing loss may begin to interfere with spatial release from informational masking.

By demonstrating a hearing-loss-independent relationship between age and spatial release from informational masking, the present study may point to declines in perceptual and/or cognitive mechanisms that play an important role in alleviating target/masker confusion. Yet, before considering such processing specific to informational masking, it is

important to consider two alternative explanations that cannot be entirely ruled out. One alternative possibility is that difficulties in localizing the masker led to poorer masking release among the older participants in the present study. This is not likely for several reasons. First, all participants verbally confirmed that they were correctly localizing the masker in each spatial condition prior to beginning the task. Second, age-group differences were similar in the F-RF and F-SUM conditions despite differences in the localization cue. This result is consistent with prior research showing similar patterns of performance in younger and older adults with “normal” hearing (≤ 25 dB HL at 250–3000 Hz) when comparing F-RF and F-SUM conditions (Li *et al.*, 2004), and conditions of physical and virtual separation (Singh *et al.*, 2008). Third, although some age-related declines in localizing based on SOA have been reported (Akeroyd and Guy, 2011; Cranford *et al.*, 1993; Cranford *et al.*, 1990; Cranford and Romereim, 1992), the characteristics of these declines do not plausibly account for the present results. For example, Cranford *et al.* (1993), Cranford *et al.* (1990), and Cranford and Romereim (1992) found that older adults committed more localization errors compared to younger adults at short SOAs within the range of summing localization, but the age-group difference was strongest at SOAs between 0.3 and 0.5 ms. Moreover, most errors were made by incorrectly localizing to the midline between the loudspeakers, rather than incorrectly localizing toward one of the loudspeakers (Cranford *et al.*, 1993), and no differences between the age groups were found at SOAs 0.7–8 ms (Cranford *et al.*, 1993; Cranford *et al.*, 1990; Cranford and Romereim, 1992). Similar to the present study, Akeroyd and Guy (2011) used a 60° separation between loudspeakers and a 4-ms SOA and found that the strength with which older adults localized speech stimuli toward the lead loudspeaker (i.e., localization dominance of the precedence effect) was variable. However, unlike the age-related effects in the present study, Akeroyd and Guy (2011) found that hearing loss influenced localization dominance, such that greater hearing loss was associated with a shift in localization away from the lead speaker toward the lag speaker. Moreover, despite a considerable range of hearing loss among the participants in the study of Akeroyd and Guy (2011), localization dominance was always strong enough to be perceived from the lead side: the shift in localization away from the lead loudspeaker was no more than 10° for the participants with “normal” hearing ($PTA_{500-4000\text{ Hz}} < 25$ dB HL), and no more than 25° for those with “mild” ($PTA_{500-4000\text{ Hz}} = 25-39$ dB HL) and “moderate” ($PTA_{500-4000\text{ Hz}} = 40-61$ dB HL) hearing loss. Given the fact that in the present study (1) age-related differences were similar in the F-RF and F-SUM conditions despite differences in the localization cue, (2) a large 60° separation was used that should have been robust to the influence of localization errors among the older adults, and (3) no independent relationship between hearing loss and masking release was found, it is unlikely that declines in localization accuracy contributed substantively to the large age-related reduction in masking release presently observed. However, such considerations do not take into account potential age-related differences in the perceived spatial

width (Whitmer *et al.*, 2012, 2013, 2014) or quality of the auditory percepts in the present study, and do not rule out potential effects that may be specific to the low-pass filtered vocoded stimuli. Therefore, future research should include objective measures of localization and spatial perception to assess their potential contributions to the age-related differences presently observed.

Another alternative explanation is that virtual spatial separation may have released the majority of informational masking for both the younger and older participants, and that the age-group differences observed in the spatially separated conditions reflect differences in baseline energetic masking. As previously discussed, the virtual separation paradigm should have minimized confounding changes in energetic masking within participants across spatial conditions. However, if energetic masking was greater to begin with for the older participants, then greater energetic masking of the target would have remained in the F-RF condition for the older participants compared to the younger participants. Such an account would require a large difference in energetic masking between the age groups that would have been present to a similar degree in all spatial conditions. Assuming that total masking comprises the sum of energetic and informational masking, this account may at first seem difficult to reconcile with the matched performance between the age groups in the F-F condition, as this would suggest that older adults experience substantially less informational masking than younger adults within challenging listening environments. However, as mentioned above, there is evidence of a ceiling effect on informational masking (Arbogast *et al.*, 2005; Freyman *et al.*, 2008), as well as evidence of other mechanisms that may limit informational masking as energetic masking is increased (see discussion in Arbogast *et al.*, 2005). Thus, based solely on the threshold data, the age-group differences may not necessarily contradict a baseline energetic masking account. However, the broader results of the present study cannot be easily reconciled with evidence that hearing loss is a main factor contributing to baseline energetic masking (e.g., Agus *et al.*, 2009; Barrenäs and Wikström, 2000; Goossens *et al.*, 2017; Humes *et al.*, 1994; Souza and Turner, 1994). Much evidence comes from speech identification tasks using energetic maskers, but Tye-Murray *et al.* (2011) found that detection thresholds of a single syllable (/ba/) in speech-shaped noise was nearly identical between younger and older adults with age-typical hearing. Similarly, research on BMLDs using target-in-noise detection tasks tend to show little difference in baseline energetic masking (i.e., NoSo threshold) between younger and older adults with age-typical or better hearing (Anderson *et al.*, 2018; Eddins and Eddins, 2018; Grose *et al.*, 1994; Novak and Anderson, 1982; Pichora-Fuller and Schneider, 1991), while higher NoSo thresholds have been observed for hearing-impaired older adults (Novak and Anderson, 1982), but not always (Eddins and Eddins, 2018). Although the present study did not directly test energetic masking, prior studies using virtual separation with energetic maskers (i.e., broadband noise) have been consistent with this broader research insofar as energetic masking effects in older adults with age-typical hearing have been relatively small (Helfer

and Freyman, 2008; Li *et al.*, 2004) and, in Helfer and Freyman (2008), correlated to some extent with hearing loss but not significantly with age. Therefore, although a baseline energetic masking account cannot be ruled out in the present study and must be tested in the future, it is not clear at the moment whether energetic masking alone could account for the entirety of the present results, given that such a large age-related decline in masking release was observed without any indication that hearing loss played a role. Thus, the present results may also point to age-related declines in perceptual and/or cognitive mechanisms thought to be involved in the spatial release from informational masking itself under challenging listening conditions.

Much remains to be known about the mechanisms underlying spatial release from informational masking, but processes involving auditory object perception and selective attention are likely to play essential roles. Indeed, across a large literature, auditory object perception and selective attention have been central concepts in the understanding of how listeners solve the cocktail party problem (Bronkhorst, 2015) and navigate complex auditory scenes more generally (Bregman, 1990), and their contributions specifically to spatial release from informational masking have been empirically supported (Ihlefeld and Shinn-Cunningham, 2008a,b). This prior research suggests that spatial separation reduces target/masker confusion by serving as a cue for segregating and maintaining the target and masker as distinct auditory objects, and for allowing attention to be better directed to the target and allowing the masker to be more easily ignored (Ihlefeld and Shinn-Cunningham, 2008b,a). Although object grouping and selective attention were not directly manipulated in the present study, age-related declines in one or both of these areas may have contributed to the results. Some studies report no significant age-related declines in the ability to fuse the lead and lag sounds into a single auditory object in the precedence effect (Lister and Roberts, 2005; Schneider *et al.*, 1994). However, Gallun *et al.* (2014) suggest that echo thresholds (SOA at which lag sound is heard as a separate source) can be higher for older compared to younger listeners. Although higher echo thresholds would mean that older adults would be *less* likely than younger adults to hear the front masker as a separate sound in the spatially separated conditions, any age-related differences in fusion could affect the perceptual quality of the masker, including its spatial width as previously mentioned (Whitmer *et al.*, 2012, 2013, 2014), in a way that may have contributed to the present results. In addition, evidence that other aspects of auditory object processing may decline with age, such as object streaming (Ben-David *et al.*, 2012; Ezzatian *et al.*, 2015) and, perhaps to some extent, object enumeration (Roberts *et al.*, 2019), suggest that age-related declines in the speed and consistency of object processing, important for detecting fleeting instances of single-syllable target words, may have contributed to poorer release from informational masking. Furthermore, there is some evidence suggesting that in the visual domain, view-invariant object recognition may decline among older adults (Burke *et al.*, 2012). If there are similar age-related declines in the ability to invariantly represent auditory objects, the older

participants may have experienced greater difficulty recognizing different targets as belonging to the same category of sound (for review, see [Heald et al., 2017](#)), although similar inter-age-group false-alarm rates in the spatially separated conditions suggests that maskers were accurately recognized as such.

In addition to early perceptual mechanisms, the ability to quickly segregate, maintain, and flexibly process auditory objects at separate locations is also likely to depend upon higher-order cognitive functions that have been shown to decline with age (for reviews, see [Anderson and Craik, 2017](#); [Drag and Bieliauskas, 2010](#)). Working memory has been linked to spatial release from masking ([Clayton et al., 2016](#)), but its influence in the present study may have been limited by the simplicity of the task and short trial length. Selective attention, on the other hand, should have been important for detecting brief targets that varied by word and onset time across trials, and is likely to play a central role in spatial release from informational masking, as evidenced across a range of studies that have manipulated attention and the speed of attentional buildup, and have tied measures of general attentional ability to listening performance under multi-talker conditions (e.g., [Best et al., 2008](#); [Best et al., 2007](#); [Clayton et al., 2016](#); [Holmes et al., 2018](#); [Ihlefeld and Shinn-Cunningham, 2008a](#); [Kidd et al., 2005](#); [Kitterick et al., 2010](#); [Oberfeld and Klöckner-Nowotny, 2016](#)). It follows that any age-related declines in the speed and control of selective attention or increased susceptibility to masker distraction (for review, see [Zanto and Gazzaley, 2014](#)) would negatively affect the ability to reduce confusion under challenging listening conditions. Thus, a hypothesis can be offered for the present results: Insofar as spatial separation provides a cue that facilitates object perception and selective attention, age-related declines in these perceptual and cognitive mechanisms may have limited the extent to which informational masking was released across spatial conditions. There has been some success using behavioral measures to tease apart the separate contributions of object perception and selective attention on spatial release from informational masking ([Ihlefeld and Shinn-Cunningham, 2008b,a](#)), but future investigations may require neurophysiological measures that can offer insight into auditory processing even when attention is directed away from the auditory domain and behavioral responses are not made. Recent event-related potential (ERP) research using a similar paradigm to the present study has shown clear indices of spatial release from informational masking that begin early in perceptual processing ([Zobel et al., 2018](#)). Future ERP research may be able to pinpoint the stages of auditory processing at which age-related declines limit the potential of a spatial cue to reduce confusion within complex, noisy environments.

The present study demonstrated a useful paradigm for studying declines in spatial release from informational masking specific to aging. More research is needed to determine exactly why age-related differences were so apparent under the present conditions compared to prior virtual separation studies (e.g., [Li et al., 2004](#)), and whether such differences will generalize to other sets of stimuli and spatial configurations. By reducing non-spatial cues and task demands,

controlling for hearing loss with low-pass filtering, matching inter-age-group performance when target and masker were spatially co-located, and measuring a large release from masking across spatial conditions, the present study may have benefited from greater power to assess declines in spatial release from informational masking. Lack of clear, robust age-specific differences in prior research using natural-speech-identification tasks may also reflect compensatory mechanisms among older adults at linguistic stages of processing that were not needed for the simple detection task used in the present study. Indeed, it is reasonable to assume that the present study did not require any speech-specific processing. After all, the ability to detect relevant signals in confusable noise, though certainly crucial for understanding speech in multi-talker environments, is not limited to speech. Likewise, spatial release from informational masking is not just useful at a cocktail party, but likely constitutes a fundamental component of how listeners generally hear and understand auditory objects within any complex acoustic environment. Findings from the present study may indicate age-related declines in perceptual and/or cognitive mechanisms underlying spatial release from informational masking that may contribute to general difficulties in navigating the complex auditory scenes of everyday life, but further investigation is required to test this hypothesis and identify the specific stages of processing at which aging may limit informational masking release.

ACKNOWLEDGMENTS

The research reported here was funded by NSF GRFP/GROW Award No. 1451512 to B.H.Z., and matching funds by University Medical Center Groningen. Further funding was provided by a VICI Grant No. 918-17-603 (awarded to D.B.) from the Netherlands Organization for Scientific Research and the Netherlands Organization for Health Research and Development. The authors thank Britt Bosma, Anne Nijman, Etienne Gaudrain, Elif Kaplan, and Mathieu Blom for their assistance in the experimental design, setup, data collection, and analysis. The authors also thank Richard L. Freyman for his helpful feedback and advice on the study and the interpretation of the results.

¹The data that support the findings of this study are openly available in DataverseNL ([Data Archiving and Networked Services, 2019](#)).

²[Qin and Oxenham \(2003\)](#) originally designed the low-pass filter cutoff criteria to capture at least some F_0 distinctions among talkers. Therefore, the vocoding process may not entirely eliminate distinguishing non-spatial cues from the stimuli, but prior research using this method suggests that for listeners, such cues among same-sex talkers are substantially reduced ([Freyman et al., 2008](#); [Morse-Fortier et al., 2017](#)).

³The equation published in [Lesmes et al. \(2015\)](#) (Eq. 12) contained a misprint. The equation expressed here [Eq. (3)], which divides the first lapse term by 2, is correct ([Lu, 2018](#)).

⁴Multicollinearity was elevated when the high-frequency hearing loss predictor was used ($VIF = 5.49$). However, results were consistent with all other reported regression models, none of which indicated a high degree of multicollinearity ($VIF \leq 2.49$).

Agus, T. R., Akeroyd, M. A., Gatehouse, S., and Warden, D. (2009). "Informational masking in young and elderly listeners for speech masked by simultaneous speech and noise," *J. Acoust. Soc. Am.* **126**(4), 1926–1940.

- Akeroyd, M. A., and Guy, F. H. (2011). "The effect of hearing impairment on localization dominance for single-word stimuli," *J. Acoust. Soc. Am.* **130**(1), 312–323.
- Anderson, N. D., and Craik, F. I. M. (2017). "50 years of cognitive aging theory," *J. Gerontol. Ser. B: Psychol. Sci. Soc. Sci.* **72**(1), 1–6.
- Anderson, S., Ellis, R., Mehta, J., and Goupell, M. J. (2018). "Age-related differences in binaural masking level differences: Behavioral and electrophysiological evidence," *J. Neurophysiol.* **120**(6), 2939–2952.
- Arbogast, T. L., Mason, C. R., and Kidd, G., Jr. (2005). "The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **117**(4), 2169–2180.
- Avivi-Reich, M., Daneman, M., and Schneider, B. A. (2014). "How age and linguistic competence alter the interplay of perceptual and cognitive factors when listening to conversations in a noisy environment," *Front. Syst. Neurosci.* **8**, 1–17.
- Barrenäs, M.-L., and Wikström, I. (2000). "The influence of hearing and age on speech recognition scores in noise in audiological patients and in the general population," *Ear Hear.* **21**(6), 569–577.
- Başkent, D., and Gaudrain, E. (2016). "Musician advantage for speech-on-speech perception," *J. Acoust. Soc. Am.* **139**(3), EL51–EL56.
- Ben-David, B. M., Tse, V. Y. Y., and Schneider, B. A. (2012). "Does it take older adults longer than younger adults to perceptually segregate a speech target from a background masker?," *Hear. Res.* **290**(1–2), 55–63.
- Bernstein, L. R., and Trahiotis, C. (2018). "Effects of interaural delay, center frequency, and no more than 'slight' hearing loss on precision of binaural processing: Empirical data and quantitative modeling," *J. Acoust. Soc. Am.* **144**(1), 292–307.
- Best, V., Ozmeral, E. J., Kopčo, N., and Shinn-Cunningham, B. G. (2008). "Object continuity enhances selective auditory attention," *Proc. Natl. Acad. Sci.* **105**(35), 13174–13178.
- Best, V., Ozmeral, E. J., and Shinn-Cunningham, B. G. (2007). "Visually-guided attention enhances target identification in a complex auditory scene," *J. Assoc. Res. Otolaryngol.* **8**(2), 294–304.
- Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, MA).
- Boersma, P., and Weenink, D. (2018). "Praat: Doing phonetics by computer (version 6.0.37) [computer program]," <http://www.praat.org/> (Last viewed June 25, 2019).
- Bradlow, A. R., and Alexander, J. A. (2007). "Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners," *J. Acoust. Soc. Am.* **121**(4), 2339–2349.
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA).
- Bronkhorst, A. W. (2000). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acta Acust. Acust.* **86**(1), 117–128.
- Bronkhorst, A. W. (2015). "The cocktail-party problem revisited: Early processing and selection of multi-talker speech," *Atten. Percept. Psychophys.* **77**(5), 1465–1487.
- Brungart, D. S. (2001). "Informational and energetic masking effects in the perception of two simultaneous talkers," *J. Acoust. Soc. Am.* **109**(3), 1101–1109.
- Brungart, D. S., Simpson, B. D., and Freyman, R. L. (2005). "Precedence-based speech segregation in a virtual auditory environment," *J. Acoust. Soc. Am.* **118**(5), 3241–3251.
- Burke, S. N., Ryan, L., and Barnes, C. A. (2012). "Characterizing cognitive aging of recognition memory and related processes in animal models and in humans," *Front. Aging Neurosci.* **4**, 1–15.
- Carhart, R., Tillman, T. W., and Greetis, E. S. (1969). "Perceptual masking in multiple sound backgrounds," *J. Acoust. Soc. Am.* **45**(3), 694–703.
- CHABA (Committee on Hearing Bioacoustics and Biomechanics) (1988). "Speech understanding and aging," *J. Acoust. Soc. Am.* **83**(3), 859–895.
- Cherry, E. C. (1953). "Some experiments on the recognition of speech, with one and with two ears," *J. Acoust. Soc. Am.* **25**(5), 975–979.
- Clark, J. G. (1981). "Uses and abuses of hearing loss classification," *ASHA* **23**(7), 493–500.
- Clayton, K. K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A. D., and Kidd, G. (2016). "Executive function, visual attention and the cocktail party problem in musicians and non-musicians," *PLoS One* **11**(7), 1–17.
- Cranford, J. L., Andres, M. A., Platz, K. K., and Reissig, K. L. (1993). "Influences of age and hearing loss on the precedence effect in sound localization," *J. Speech Lang. Hear. Res.* **36**, 437–441.
- Cranford, J. L., Boose, M., and Moore, C. A. (1990). "Effects of aging on the precedence effect in sound localization," *J. Speech Lang. Hear. Res.* **33**, 654–659.
- Cranford, J. L., and Romereim, B. (1992). "Precedence effect and speech understanding in elderly listeners," *J. Am. Acad. Audiol.* **3**(6), 405–409.
- Culling, J. F., and Summerfield, Q. (1995). "The role of frequency modulation in the perceptual segregation of concurrent vowels," *J. Acoust. Soc. Am.* **98**(2), 837–846.
- Darwin, C. J., Brungart, D. S., and Simpson, B. D. (2003). "Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers," *J. Acoust. Soc. Am.* **114**(5), 2913–2922.
- Darwin, C. J., and Hukin, R. W. (2000). "Effectiveness of spatial cues, prosody, and talker characteristics in selective attention," *J. Acoust. Soc. Am.* **107**(2), 970–977.
- Data Archiving and Networked Services (2019). <https://hdl.handle.net/10411/Z4VUQU> (Last viewed June 25, 2019).
- Drag, L. L., and Bieliauskas, L. A. (2010). "Contemporary review 2009: Cognitive aging," *J. Geriatr. Psych. Neurol.* **23**(2), 75–93.
- Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., and Kidd, G., Jr. (2003). "Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity," *J. Acoust. Soc. Am.* **114**(1), 368–379.
- Eddins, A. C., and Eddins, D. A. (2018). "Cortical correlates of binaural temporal processing deficits in older adults," *Ear Hear.* **39**(3), 594–604.
- Eddins, A. C., Ozmeral, E. J., and Eddins, D. A. (2018). "How aging impacts the encoding of binaural cues and the perception of auditory space," *Hear. Res.* **369**, 79–89.
- El Boghdady, N., Gaudrain, E., and Başkent, D. (2019). "Does good perception of vocal characteristics relate to better speech-on-speech intelligibility for cochlear implant users?," *J. Acoust. Soc. Am.* **145**(1), 417–439.
- Ezzatian, P., Li, L., Pichora-Fuller, K., and Schneider, B. A. (2015). "Delayed stream segregation in older adults: More than just informational masking," *Ear Hear.* **36**(4), 482–484.
- Fletcher, H. (1940). "Auditory patterns," *Rev. Mod. Phys.* **12**(12), 47–65.
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state: A practical method for grading the cognitive state of patients for the clinician," *J. Psych. Res.* **12**(3), 189–198.
- Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2001). "Spatial release from informational masking in speech recognition," *J. Acoust. Soc. Am.* **109**(5), 2112–2122.
- Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2004). "Effect of number of masking talkers and auditory priming on informational masking in speech recognition," *J. Acoust. Soc. Am.* **115**(5), 2246–2256.
- Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2008). "Spatial release from masking with noise-vocoded speech," *J. Acoust. Soc. Am.* **124**(3), 1627–1637.
- Freyman, R. L., Helfer, K. S., and Balakrishnan, U. (2005). "Spatial and spectral factors in release from informational masking in speech recognition," *Acta Acust. Acust.* **91**(3), 537–545.
- Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (1999). "The role of perceived spatial separation in the unmasking of speech," *J. Acoust. Soc. Am.* **106**(6), 3578–3588.
- Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2015). "Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition," *Front. Aging Neurosci.* **6**, 1–25.
- Gallun, F. J., Diedesch, A. C., Kappel, S. D., and Jakien, K. M. (2013). "Independent impacts of age and hearing loss on spatial release in a complex auditory environment," *Front. Neurosci.* **7**, 252.
- Gallun, F. J., McMillan, G. P., Molis, M. R., Kappel, S. D., Dann, S. M., and Konrad-Martin, D. L. (2014). "Relating age and hearing loss to monaural, bilateral, and binaural temporal sensitivity," *Front. Neurosci.* **8**, 172.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**(1), 103–138.
- Glyde, H., Cameron, S., Dillon, H., Hickson, L., and Seeto, M. (2013). "The effects of hearing impairment and aging on spatial processing," *Ear Hear.* **34**(1), 15–28.
- Glyde, H., Hickson, L., Cameron, S., and Dillon, H. (2011). "Problems hearing in noise in older adults: A review of spatial processing disorder," *Trends Amplif.* **15**(3), 116–126.
- Goossens, T., Vercammen, C., Wouters, J., and van Wieringen, A. (2017). "Masked speech perception across the adult lifespan: Impact of age and hearing impairment," *Hear. Res.* **344**, 109–124.
- Gordon-Salant, S. (2005). "Hearing loss and aging: New research findings and clinical implications," *J. Rehab. Res. Dev.* **42**, 9–23.
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, Oxford).

- Grose, J. H., Poth, E. A., and Peters, R. W. (1994). "Masking level differences for tones and speech in elderly listeners with relatively normal audiograms," *J. Speech Lang. Hear. Res.* **37**(2), 422–428.
- Hannula, S., Bloigu, R., Majamaa, K., Sorri, M., and Mäki-Torkko, E. (2011). "Audiogram configurations among older adults: Prevalence and relation to self-reported hearing problems," *Int. J. Audiol.* **50**(11), 793–801.
- Heald, S. L. M., Van Hedger, S. C., and Nusbaum, H. C. (2017). "Perceptual plasticity for auditory object recognition," *Front. Psych.* **8**, 1–16.
- Helfer, K. S., Chevalier, J., and Freyman, R. L. (2010). "Aging, spatial cues, and single- versus dual-task performance in competing speech perception," *J. Acoust. Soc. Am.* **128**(6), 3625–3633.
- Helfer, K. S., and Freyman, R. L. (2008). "Aging and speech-on-speech masking," *Ear Hear.* **29**(1), 87–98.
- Hirsh, I. J. (1948). "The influence of interaural phase on interaural summation and inhibition," *J. Acoust. Soc. Am.* **20**(4), 536–544.
- Hoffman, H. J., Dobie, R. A., Ko, C.-W., Themann, C. L., and Murphy, W. J. (2012). "Hearing threshold levels at age 70 years (65–74 years) in the unscreened older adult population of the United States, 1959–1962 and 1999–2006," *Ear Hear.* **33**(3), 437–440.
- Holmes, E., Kitterick, P. T., and Summerfield, A. Q. (2018). "Cueing listeners to attend to a target talker progressively improves word report as the duration of the cue-target interval lengthens to 2,000 ms," *Atten. Percept. Psychophys.* **80**(6), 1520–1538.
- Humes, L. E., Watson, B. U., Christensen, L. A., Cokely, C. G., Halling, D. C., and Lee, L. (1994). "Factors associated with individual differences in clinical measures of speech recognition among the elderly," *J. Speech Lang. Hear. Res.* **37**(2), 465–474.
- Hummersone, C. (2017). "The IoSR MATLAB toolbox, (version 2.8) [computer software]," Institute of Sound Recording, University of Surrey Guildford, Surrey, UK.
- IBM (2015). "SPSS statistics for Macintosh, (version 23) [computer software]," Armonk, NY.
- Ilhfeld, A., and Shinn-Cunningham, B. (2008a). "Disentangling the effects of spatial cues on selection and formation of auditory objects," *J. Acoust. Soc. Am.* **124**(4), 2224–2235.
- Ilhfeld, A., and Shinn-Cunningham, B. (2008b). "Spatial release from energetic and informational masking in a selective speech identification task," *J. Acoust. Soc. Am.* **123**(6), 4369–4379.
- ISO (2010). ISO 8253-1:2010: *Acoustics—Audiometric Test Methods—Part 1: Pure-tone Air and Bone Conduction Audiometry* (International Organization for Standardization, Geneva, Switzerland).
- ISO (2017). ISO 7029:2017: *Acoustics—Statistical Distribution of Hearing Thresholds Related to Age and Gender* (International Organization for Standardization, Geneva, Switzerland).
- Jakien, K. M., and Gallun, F. J. (2018). "Normative data for a rapid, automated test of spatial release from masking," *Am. J. Audiol.* **27**, 529–538.
- Jakien, K. M., Kampel, S. D., Gordon, S. Y., and Gallun, F. J. (2017). "The benefits of increased sensation level and bandwidth for spatial release from masking," *Ear Hear.* **38**(1), e13–e21.
- Kaernbach, C. (1990). "A single-interval adjustment-matrix (SIAM) procedure for unbiased adaptive testing," *J. Acoust. Soc. Am.* **88**(6), 2645–2655.
- Kidd, G., Arbogast, T. L., Mason, C. R., and Gallun, F. J. (2005). "The advantage of knowing where to listen," *J. Acoust. Soc. Am.* **118**(6), 3804–3815.
- Kidd, G., Mason, C. R., Best, V., and Marrone, N. (2010). "Stimulus factors influencing spatial release from speech-on-speech masking," *J. Acoust. Soc. Am.* **128**(4), 1965–1978.
- Kidd, G., Mason, C. R., Deliwal, P. S., Woods, W. S., and Colburn, H. S. (1994). "Reducing informational masking by sound segregation," *J. Acoust. Soc. Am.* **95**(6), 3475–3480.
- Kidd, G., Jr., Mason, C. R., Richards, V. M., Gallun, F. J., and Durlach, N. I. (2008). "Informational masking," in *Auditory Perception of Sound Sources*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer Science+Business Media, LLC, New York), pp. 143–189.
- Kitterick, P. T., Bailey, P. J., and Summerfield, A. Q. (2010). "Benefits of knowing who, where, and when in multi-talker listening," *J. Acoust. Soc. Am.* **127**(4), 2498–2508.
- Klein, S. A. (2001). "Measuring, estimating, and understanding the psychometric function: A commentary," *Percept. Psychophys.* **63**(8), 1421–1455.
- Lesmes, L. A., Lu, Z.-L., Baek, J., Tran, N., Doshier, B. A., and Albright, T. D. (2015). "Developing Bayesian adaptive methods for estimating sensitivity thresholds (d') in Yes-No and forced-choice tasks," *Front. Psych.* **6**, 1–24.
- Li, L., Daneman, M., Qi, J. G., and Schneider, B. A. (2004). "Does the information content of an irrelevant source differentially affect spoken word recognition in younger and older adults?," *J. Exp. Psych.: Human Percept. Perform.* **30**(6), 1077–1091.
- Licklider, J. C. R. (1948). "The influence of interaural phase relations upon the masking of speech by white noise," *J. Acoust. Soc. Am.* **20**(2), 150–159.
- Lister, J. J., and Roberts, R. A. (2005). "Effects of age and hearing loss on gap detection and the precedence effect: Narrow-band stimuli," *J. Speech Lang. Hear. Res.* **48**(2), 482–493.
- Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (1999). "The precedence effect," *J. Acoust. Soc. Am.* **106**, 1633–1654.
- Lu, Z. (2018). (private communication).
- Lutfi, R. A. (1993). "A model of auditory pattern analysis based on component-relative-entropy," *J. Acoust. Soc. Am.* **94**(2), 748–758.
- Lutfi, R. A., Gilbertson, L., Heo, I., Chang, A.-C., and Stamas, J. (2013). "The information-divergence hypothesis of informational masking," *J. Acoust. Soc. Am.* **134**(3), 2160–2170.
- Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*, 2nd ed. (Lawrence Erlbaum Associates, Mahwah, NJ).
- MathWorks Inc. (2015). "MATLAB, (version R2015a) [computer software]," Natick, MA.
- Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). "Speech recognition in adverse conditions: A review," *Lang. Cogn. Process.* **27**(7-8), 953–978.
- Miller, G. A. (1947). "The masking of speech," *Psych. Bull.* **44**(2), 105–129.
- Morse-Fortier, C., Parrish, M. M., Baran, J. A., and Freyman, R. L. (2017). "The effects of musical training on speech detection in the presence of informational and energetic masking," *Trends Hear.* **21**, 1–12.
- Neff, D. L. (1995). "Signal properties that reduce masking by simultaneous, random-frequency maskers," *J. Acoust. Soc. Am.* **98**(4), 1909–1920.
- Neff, D. L., and Green, D. M. (1987). "Masking produced by spectral uncertainty with multicomponent maskers," *Percept. Psychophys.* **41**(5), 409–415.
- Novak, R. E., and Anderson, C. V. (1982). "Differentiation of types of presbycusis using the masking-level difference," *J. Speech Lang. Hear. Res.* **25**(4), 504–508.
- Oberfeld, D., and Klöckner-Nowotny, F. (2016). "Individual differences in selective attention predict speech identification at a cocktail party," *ELife* **5**, 1–24.
- Oh, E. L., and Lutfi, R. A. (2000). "Effect of masker harmonicity on informational masking," *J. Acoust. Soc. Am.* **108**(2), 706–709.
- Pichora-Fuller, M. K., Alain, C., and Schneider, B. A. (2017). "Older adults at the cocktail party," in *The Auditory System at the Cocktail Party*, edited by J. C. Middlebrooks, J. Z. Simon, A. N. Popper, and R. R. Fay (Springer International, Cham), Vol. 60, pp. 227–259.
- Pichora-Fuller, M. K., and Schneider, B. A. (1991). "Masking-level differences in the elderly: A comparison of antiphasic and time-delay dichotic conditions," *J. Speech Lang. Hear. Res.* **34**(6), 1410–1422.
- Plomp, R., and Mimpen, A. M. (1979). "Improving the reliability of testing the speech reception threshold for sentences," *Audiology* **18**(1), 43–52.
- Preacher, K. J., Rucker, D. D., MacCallum, R. C., and Nicewander, W. A. (2005). "Use of the extreme groups approach: A critical reexamination and new recommendations," *Psych. Methods* **10**(2), 178–192.
- Psychology Software Tools, Inc. (2016). "E-Prime, (version 3.0) [computer software]," Pittsburgh, PA.
- Qin, M. K., and Oxenham, A. J. (2003). "Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers," *J. Acoust. Soc. Am.* **114**(1), 446–454.
- Rakerd, B., Aaronson, N. L., and Hartmann, W. M. (2006). "Release from speech-on-speech masking by adding a delayed masker at a different location," *J. Acoust. Soc. Am.* **119**(3), 1597–1605.
- Roberts, K. L., Doherty, N. J., Maylor, E. A., and Watson, D. G. (2019). "Can auditory objects be subitized?," *J. Exp. Psych.: Human Percept. Perform.* **45**(1), 1–15.
- Saliba, I., Bergeron, M., Martineau, G., and Chagnon, M. (2011). "Rule 3,000: A more reliable precursor to perceive vestibular schwannoma on MRI in screened asymmetric sensorineural hearing loss," *Eur. Arch. Oto-Rhino-Laryngol.* **268**(2), 207–212.
- Schneider, B. A., Pichora-Fuller, M. K., Kowalchuk, D., and Lamb, M. (1994). "Gap detection and the precedence effect in young and old adults," *J. Acoust. Soc. Am.* **95**(2), 980–991.

- Shaw, E. A. G. (1974). "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," *J. Acoust. Soc. Am.* **56**(6), 1848–1861.
- Singh, G., Pichora-Fuller, M. K., and Schneider, B. A. (2008). "The effect of age on auditory spatial attention in conditions of real and simulated spatial separation," *J. Acoust. Soc. Am.* **124**(2), 1294–1305.
- Souza, P. E., and Turner, C. W. (1994). "Masking of speech in young and elderly listeners with hearing loss," *J. Speech Lang. Hear. Res.* **37**, 655–661.
- Srinivasan, N. K., Jakien, K. M., and Gallun, F. J. (2016). "Release from masking for small spatial separations: Effects of age and hearing loss," *J. Acoust. Soc. Am.* **140**(1), EL73–EL78.
- Srinivasan, N. K., Stansell, M., and Gallun, F. J. (2017). "The role of early and late reflections on spatial release from masking: Effects of age and hearing loss," *J. Acoust. Soc. Am.* **141**(3), EL185–EL191.
- Stanislaw, H., and Todorov, N. (1999). "Calculation of signal detection theory measures," *Behav. Res. Meth. Instrum. Comput.* **31**(1), 137–149.
- Tye-Murray, N., Spehar, B., Myerson, J., Sommers, M. S., and Hale, S. (2011). "Crossmodal enhancement of speech detection in young and older adults: Does signal content matter?," *Ear Hear.* **32**(5), 650–655.
- Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000). "Method for the selection of sentence materials for efficient measurement of the speech reception threshold," *J. Acoust. Soc. Am.* **107**(3), 1671–1684.
- Vestergaard, M. D., Fyson, N. R. C., and Patterson, R. D. (2009). "The interaction of vocal characteristics and audibility in the recognition of concurrent syllables," *J. Acoust. Soc. Am.* **125**(2), 1114–1124.
- Watson, C. S., Kelly, W. J., and Wroton, H. W. (1976). "Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty," *J. Acoust. Soc. Am.* **60**(5), 1176–1186.
- Watson, C. S., Wroton, H. W., Kelly, W. J., and Benbassat, C. A. (1975). "Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent intervals," *J. Acoust. Soc. Am.* **57**(5), 1175–1185.
- Whitmer, W. M., Seeber, B. U., and Akeroyd, M. A. (2012). "Apparent auditory source width insensitivity in older hearing-impaired individuals," *J. Acoust. Soc. Am.* **132**(1), 369–379.
- Whitmer, W. M., Seeber, B. U., and Akeroyd, M. A. (2013). "Measuring the apparent width of auditory sources in normal and impaired hearing," *Adv. Exp. Med. Biol.* **787**, 303–310.
- Whitmer, W. M., Seeber, B. U., and Akeroyd, M. A. (2014). "The perception of apparent auditory source width in hearing-impaired adults," *J. Acoust. Soc. Am.* **135**(6), 3548–3559.
- Wichmann, F. A., and Hill, N. J. (2001). "The psychometric function: I. Fitting, sampling, and goodness of fit," *Percept. Psychophys.* **63**(8), 1293–1313.
- Zanto, T. P., and Gazzaley, A. (2014). "Attention and ageing," in *The Oxford Handbook of Attention*, edited by A. C. (Kia) Nobre and S. Kastner (Oxford University Press, Oxford). Vol. 1, pp. 927–971.
- Zobel, B. H., Sanders, L. D., and Freyman, R. L. (2018). "Preattentive processing in the spatial unmasking of speech," poster presented at the *10th Annual Speech in Noise Workshop*, Glasgow, UK.
- Zurek, P. M. (1993). "Binaural advantages and directional effects in speech intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, 2nd ed., edited by G. A. Studebaker and I. Hochberg (Allyn and Bacon, Boston, MA), pp. 255–276.